

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2003, 08:53:35 ; Search time 4682 Seconds  
(without alignments)  
8160.955 Million cell updates/sec

Title: US-10-030-613-3

Perfect score: 934

Sequence: 1 agaatcattcttcgcggaa.....gtataaccattatttagaa 934

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_wi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	934	100.0	934	6	AX074351	AX074351 Sequence
2	934	100.0	3159	6	AX420425	AX420425 Sequence
3	934	100.0	3261	9	AK090990	AK090990 Homo sapi
4	930.8	99.7	3240	9	AY034085	AY034085 Homo sapi
5	930.8	99.7	3253	9	AY034086	AY034086 Homo sapi
6	720.8	77.2	2523	9	BC038509	BC038509 Homo sapi
7	720.8	77.2	3184	6	AX329596	AX329596 Sequence
8	720.8	77.2	3184	6	AX333015	AX333015 Sequence
9	720.8	77.2	3184	6	AX365337	AX365337 Sequence
10	720.8	77.2	3184	6	AX420436	AX420436 Sequence
11	720.8	77.2	3184	9	HUMZAK14	D83407 ZAK1-4 mRNA
12	685	73.3	3295	10	AB061525	AB061525 Mus muscu
13	666.2	71.3	953	10	AF459023	AF459023 Rattus no
14	514.6	55.1	3080	10	BC047153	BC047153 Mus muscu
15	514.6	55.1	3263	10	BC049096	BC049096 Mus muscu
16	496.2	53.1	833	10	AF459022	AF459022 Rattus no
17	457.6	49.0	3227	10	AB061524	AB061524 Mus muscu
18	422	45.2	1021	6	AX420427	AX420427 Sequence
19	420	45.0	659	10	AF237887	AF237887 Mus muscu
20	413.2	44.2	594	6	AX365318	AX365318 Sequence
21	413.2	44.2	594	10	AF237791	AF237791 Mus muscu
22	374	40.0	128461	9	AL390741	AL390741 Human DNA
23	363	38.9	134443	2	AC015547	AC015547 Homo sapi
24	280.8	30.1	2827	9	BC035854	BC035854 Homo sapi
25	279.2	29.9	828	6	AX365330	AX365330 Sequence
26	279.2	29.9	828	6	AX420437	AX420437 Sequence
27	279.2	29.9	828	9	AF176116	AF176116 Homo sapi
28	268	28.7	776	10	AF237888	AF237888 Mus muscu
29	257.6	27.6	615	6	AX420432	AX420432 Sequence
30	255	27.3	226218	2	AC112622	AC112622 Rattus no
31	255	27.3	246030	2	AC126591	AC126591 Rattus no
32	231.6	24.8	720	6	AX365333	AX365333 Sequence
33	231.6	24.8	720	9	AF176117	AF176117 Homo sapi
34	230.2	24.6	184672	2	AC107850	AC107850 Mus muscu
35	221.6	23.7	169841	2	AC073888	AC073888 Homo sapi
36	221.6	23.7	199200	9	AL359633	AL359633 Human DNA
37	219.2	23.5	67395	2	AC134417	AC134417 Mus muscu
38	217	23.2	626	10	AB075973	AB075973 Rattus no
39	214.8	23.0	2216	10	CG060263	CG060263 Cricetus
40	212.2	22.7	621	10	AF282255	AF282255 Mus muscu
41	212.2	22.7	2125	10	AF260717	AF260717 Mus muscu
42	211	22.6	597	6	AX365312	AX365312 Sequence
43	211	22.6	597	10	AF237789	AF237789 Mus muscu
44	210.8	22.6	2198	10	BC013551	BC013551 Mus muscu
45	209.6	22.4	597	6	AX365315	AX365315 Sequence

# ALIGNMENTS

RESULT 1	AX074351	Sequence 3 from Patent WO0104305.	934 bp	DNA	linear	PAT 06-FEB-2001
LOCUS	AX074351					
DEFINITION	Sequence 3 from Patent WO0104305.					
ACCESSION	AX074351					
VERSION	AX074351.1	GI:12710514				
KEYWORDS						
SOURCE						
ORGANISM	Homo sapiens (human)					
REFERENCE						
AUTHORS	Homo sapiens					
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
	Tang, Y.T. and Yue, H.					
	Human proteins involved in detoxification					
	Patent: WO 0104305-A 3 18-JAN-2001;					

Pred. No. is the number of results predicted by chance to have a

	Query Match	100.0%;	Score 934;	DB 6;	Length 3159;
	Best Local Similarity	100.0%;	Pred. No. 1.6e-285;	Mismatches 0;	Gaps 0;
	Matches 934;	Conservative 0;			
QY	1	AGAATCATACCTTCATCGGAATGAGGAGCCCGCAGGGCAGCAGGACAGCTCCCTGAGATGG	60		
Db	229	AGAATCATACCTTCATCGGAATGAGGAGCCCGCAGGGCAGCAGGACAGCTCCCTGAGATGG	288		
QY	61	AGGACTTTTCTACTGTGCTGCATAGACAGGGACTGGGCTGTCACTCTGTTGTTTGTCAGA	120		
Db	289	AGGACTTTTCTACTGTGCTGCATAGACAGGGACTGGGCTGTCACTCTGTTGTTTGTCAGA	348		
QY	121	AGAAAGCCTTTCGAACAATCACTGACTTCATGACCTCCCCAAGCTTCGTTGTTTCGCTGCAA	180		
Db	349	AGAAAGCCTTTCGAACAATCACTGACTTCATGACCTCCCCAAGCTTCGTTGTTTCGCTGCAA	408		
QY	181	TGTTCCACCAGTCAGTGTGTTGAAGGAGAGAGACGAAGAAAATTTTCAGGGACTGTTTCG	240		
Db	409	TGTTCCACCAGTCAGTGTGTTGAAGGAGAGAGACGAAGAAAATTTTCAGGGACTGTTTCG	468		
QY	241	GACTTATGATGACTGTGTGACGCTCCAGCTATTATAAGAGTTTCAGAGAGTGTCCGTATAAA	300		
Db	469	GACTTATGATGACTGTGTGACGCTCCAGCTATTATAAGAGTTTCAGAGAGTGTCCGTATAAA	528		
QY	301	CTTCAGCAATCCCTAAATCTGCAGCCCGAGCTAGGATAGAGTTCATGAACCCCAATTCAG	360		
Db	529	CTTCAGCAATCCCTAAATCTGCAGCCCGAGCTAGGATAGAGTTCATGAACCCCAATTCAG	588		
QY	361	AGGAAAAAATTAAGAGCTCTACTTTTCACAGAGTTTCAGACTCCAGAGACAGATGAGAGCAA	420		
Db	589	AGGAAAAAATTAAGAGCTCTACTTTTCACAGAGTTTCAGACTCCAGAGACAGATGAGAGCAA	648		
QY	421	ACTGCATTTGGCTCCAGCCCAGGCTGCCAAACAGTTTTCTCATCTCGCCCCCTTCCTCCCC	480		
Db	649	ACTGCATTTGGCTCCAGCCCAGGCTGCCAAACAGTTTTCTCATCTCGCCCCCTTCCTCCCC	708		
QY	481	ACCTGTTGGCTGGCAGCCCATCAACGATGCCAGCCAGTCTCTCAACTATGACCTCCTCTA	540		
Db	709	ACCTGTTGGCTGGCAGCCCATCAACGATGCCAGCCAGTCTCTCAACTATGACCTCCTCTA	768		
QY	541	TGCTGTGCCCCAACCTTAGCACAGGAGAGTATGAGCTCCATGCGAGGAGCTGAGTCCAC	600		
Db	769	TGCTGTGCCCCAACCTTAGCACAGGAGAGTATGAGCTCCATGCGAGGAGCTGAGTCCAC	828		



DEFINITION Homo sapiens calcineurin inhibitor ZAKI-4 beta splice variant 1  
 mRNA, complete cds, alternatively spliced.

ACCESSION AY034085  
 VERSION AY034085.1 GI:21307622

KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3240)  
 Cao,X., Kamebe,F., Miyazaki,T., Sarkar,D., Ohmori,S. and Seo,H.  
 Novel human ZAKI-4 isoforms: hormonal and tissue-specific  
 regulation and function as calcineurin inhibitors  
 Biochem. J. 367 (Pt 2), 459-466 (2002)

JOURNAL 22250686  
 MEDLINE 12102656  
 PUBMED 12102656

REFERENCE 2 (bases 1 to 3240)  
 Cao,X., Kamebe,F., Miyazaki,T., Ohmori,S. and Seo,H.  
 Direct Submission (08-MAY-2001) Endocrinology and Metabolism, Research  
 Submitted (08-MAY-2001) Endocrinology and Metabolism, Research  
 Institute of Environmental Medicine, Nagoya University, Furo-cho,  
 Nagoya 464-8601, Japan  
 Location/Qualifiers  
 1. 3240  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /map="6p12"  
 /tissue\_type="brain"  
 108. 839  
 /note="alternatively spliced"  
 /codon\_start=1  
 /product="calcineurin inhibitor ZAKI-4 beta splice variant  
 1"

CDS  
 1  
 /protein\_id="AAK59805.1"  
 /db\_xref="GI:21307623"  
 /translation="MRGESVFIQMRSEVFGHVPDGGFLFLLCCIDRWAVTRCAEE  
 AFQATIDNDI.PNSL.FACNVHVSFEGESEKFGGLFRTDYDCVTFQLFKFSRVRI  
 NFSPKSAARARIELHETQFGKKLYFAQVOTPETDGDGDLHLAPPQAKOFLLSPP  
 SSPVSWQFINDATPVLNLDLLYAVAKLGPGEKVELHAGTSTPSVVVHVHVDSDIEEE  
 EPKTPSPKPKLIQRGLPFSVN"  
 333. 3240  
 /note="similar to sequence with GenBank Accession Number  
 D83407"

misc\_feature 1079. 1084  
 /note="alternative"

polyA\_signal 3224. 3229  
 /note="alternative"

polyA\_signal 947 a 688 c 677 g 928 t

BASE COUNT 99.7% Score 930.8; DB 9; Length 3240;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-284;  
 Matches 932; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN  
 1 AGAATCATCTTCATCGGAATGAGGAGCCAGGCGAGGACACGCTCCCTGAAGATGG 60  
 116 AGAATCATCTTCATCGGAATGAGGAGCCAGGCGAGGACACGCTCCCTGAAGATGG 175  
 61 AGGACTTTCTTACTGTGCTGCATAGACAGGACTGGCGTGTGCACTCGTTGTTTGCAGA 120  
 116 AGAATCATCTTCATCGGAATGAGGAGCCAGGCGAGGACTGGCGTGTGCACTCGTTGTTTGCAGA 235  
 176 AGGACTTTCTTACTGTGCTGCATAGACAGGACTGGCGTGTGCACTCGTTGTTTGCAGA 295  
 121 AGAAGCTTTCAAGCAATCACTGACTTCAATGACTCCCAACTCGTTGTTTGCAGCA 180  
 236 AGAAGCTTTCAAGCAATCACTGACTTCAATGACTCCCAACTCGTTGTTTGCAGCA 295  
 181 TGTTACCAAGTCAAGTCTTTGAAGGAGAGAGAGCAAGAAAAAATTGAGGACTGTTTCG 240  
 296 TGTTACCAAGTCAAGTCTTTGAAGGAGAGAGAGCAAGAAAAAATTGAGGACTGTTTCG 355  
 241 GACTTATGATGACTGTGTGACGTTCCAGCTATTAAAGAGATTTCAGAGCTGTCCGATATA 355





BASE COUNT		755 a	559 c	501 g	708 t
ORIGIN					
Query Match	77.2%	Score 720.8;	DB 9;	Length 2523;	
Best Local Similarity	95.9%;	Pred. No. 1.3e-217;			
Matches 740;	Conservative 0;	Mismatches 32;	Indels 0;	Gaps 0;	
QY	163	CTCGTGTGTTGGCGTCAATGTTCCACAGTCAGTCTGTTTGAAGGACAAGAGACGAAGAA	222		
DB	229	CACCTCTGGTGTGCTGTGGTGGATGTCGAGGCTCTTTACCAATCAGGAGGTTAAGAA	288		
QY	223	ATTTGAGGAGACTGTTTCGGAGCTTATGATGACTGTGTGACGTTTCAGCTATTATTAAGATTT	282		
DB	289	ATTTGAGGAGACTGTTTCGGAGCTTATGATGACTGTGTGACGTTTCAGCTATTATTAAGATTT	348		
QY	283	CAGACGTGTCGATATAAACTTCAGCAATCCTAAATCTGCAGCCCGAGCTAGGATAGAGCT	342		
DB	349	CAGACGTGTCGATATAAACTTCAGCAATCCTAAATCTGCAGCCCGAGCTAGGATAGAGCT	408		
QY	343	TCATGAACCCAAATTCAGAGGAAAAAATAAAGCTCTACTTTGCACAGGTTTCAGACTCC	402		
DB	409	TCATGAACCCAAATTCAGAGGAAAAAATAAAGCTCTACTTTGCACAGGTTTCAGACTCC	468		
QY	403	AGACACAGATGGAGACAAACTGCACATGGCTCCACCCAGCCCTGCCAAACAGTTTCTCAT	462		
DB	469	AGACACAGATGGAGACAAACTGCACATGGCTCCACCCAGCCCTGCCAAACAGTTTCTCAT	528		
QY	463	CTCGCCCCCTTCCTCCCCACCTGTTGCTGGCAGCCCATCAACGATGCCACGCCAGTCTCT	522		
DB	529	CTCGCCCCCTTCCTCCCCACCTGTTGCTGGCAGCCCATCAACGATGCCACGCCAGTCTCT	588		
QY	523	CAACTATGACCTTCCTATGCTGTGGCCAAACTAGGACGAGGAGAGAGATGAGCTCCA	582		
DB	589	CAACTATGACCTTCCTATGCTGTGGCCAAACTAGGACGAGGAGAGAGATGAGCTCCA	648		
QY	583	TCGAGGAGCTAGTCCACCCCAAGTGTGCTGTCGACAGTGTGCGACAGTGCACATAGAGGA	642		
DB	649	TCGAGGAGCTAGTCCACCCCAAGTGTGCTGTCGACAGTGTGCGACAGTGCACATAGAGGA	708		
QY	643	AGAAGAGGACCCAAAGACTTCCCAAAAGCAAAATCATCAAACTCGCGCTCTCTGGCCT	702		
DB	709	AGAAGAGGACCCAAAGACTTCCCAAAAGCAAAATCATCAAACTCGCGCTCTCTGGCCT	768		
QY	703	GCCACCCCTCGTGTCCAACTGAGCTGCCTTCCTTCGATATAAGCCGCTCTCTCTTT	762		
DB	769	GCCACCCCTCGTGTCCAACTGAGCTGCCTTCCTTCGATATAAGCCGCTCTCTCTTT	828		
QY	763	ATCATGCTTTTCCCTGTTGTTCTCAAAAAAATTCCTTTAAATTCCTGGGTGTTT	822		
DB	829	ATCATGCTTTTCCCTGTTGTTCTCAAAAAAATTCCTTTAAATTCCTGGGTGTTT	888		
QY	823	GGTTGTTTGAGATTTCCTTCCTGTTATCAAGCCTCTCGACAAAAGGGCTAGCAAAAGCT	882		
DB	889	GGTTGTTTGAGATTTCCTTCCTGTTATCAAGCCTCTCGACAAAAGGGCTAGCAAAAGCT	948		
QY	883	GATATGCTCTCTGATCATATCATACCCATTAAAGTATATAACCCATTATTTAGAA	934		
DB	949	GATATGCTCTCTGATCATATCATACCCATTAAAGTATATAACCCATTATTTAGAA	1000		

<b>RESULT 7</b>			
AX329596	AX329596	3184 bp	DNA
LOCUS	Sequence 105 from Patent WO0194629.		
DEFINITION			linear
			PAT 09-JAN-2002

Accession	AX329596	GI:18102574
Version	AX329596.1	
Keywords		
Source	Homo sapiens (human)	
Organism	Homo sapiens	
Reference	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
Authors	Young P.E., Augustus M., Carter K.C., Ebner R., Endress G.,	
Title	Horrigan S., Soppet D.R. and Weaver Z.	
Journal	Cancer gene determination and therapeutic screening using signature gene sets	
Patent	WO 0194629-A 105 13-DEC-2001;	
Journal	Avalon Pharmaceuticals (US)	
Features	Location/Qualifiers	
Source	1..3184	
Base Count	921 a 681 c 657 g 925 t	
Origin		
Query Match	77.2%; Score 720.8; DB 6; Length 3184;	
Best Local Similarity	95.9%; Pred. No. 1.4e-217;	
Matches	740; Conservative 0; Mismatches 32; Indels 0; Gaps 0;	
163	CTCGTTGTTGGTGCAATGTTCCACGCTCAGTGTTTGAAGAGAAGAGCAGAGAA 222	
164		
165		
166		
167		
168		
169		
170		
171		
172		
173		
174		
175		
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191		
192		
193		
194		
195		
196		
197		
198		
199		
200		
201		
202		
203		
204		
205		
206		
207		
208		
209		
210		
211		
212		
213		
214		
215		
216		
217		
218		
219		
220		
221		
222		
223		
224		
225		
226		
227		
228		
229		
230		
231		
232		
233		
234		
235		
236		
237		
238		
239		
240		
241		
242		
243		
244		
245		
246		
247		
248		
249		
250		
251		
252		
253		
254		
255		
256		
257		
258		
259		
260		
261		
262		
263		
264		
265		
266		
267		
268		
269		
270		
271		
272		
273		
274		
275		
276		
277		
278		
279		
280		
281		
282		
283		
284		
285		
286		
287		
288		
289		
290		
291		
292		
293		
294		
295		
296		
297		
298		
299		

```

QY 883 GATATGTCCTCTCATATCATATACCCATTAGTATACCCATTATTTAGAA 934
Db 942 GATATGTCCTCTCATATCATATACCCATTAGTATACCCATTATTTAGAA 993

RESULT 8
AX333015
LOCUS 3184 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 3524 from Patent WO0194629.
ACCESSION AX333015
VERSION AX333015.1 GI:18123649
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R., and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 3524 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
Location/Qualifiers
1..3184
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 921 a 681 c 657 g 925 t
ORIGIN
Query Match 77.2%; Score 720.8; DB 6; Length 3184;
Best Local Similarity 95.9%; Pred. No. 1.4e-217;
Matches 740; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 163 CTCGTCGTCCTGTCGTCATGTCACAGTCAGTGTGTTGAAGGAGAGAGAGAGAA 222
Db 222 CACTCTGTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 281
QY 223 ATTTGAGGAGCTGTTTCGGAGCTTATGATGACTGTGTGACGTTCCAGCTATTTAAGAGTTT 282
Db 282 ATTTGAGGAGCTGTTTCGGAGCTTATGATGACTGTGTGACGTTCCAGCTATTTAAGAGTTT 341
QY 283 CAGACGTCGTCGTCATTAACCTCAGCAATCCTAAATCTGACGCCGAGCTAGGATAGCT 342
Db 342 CAGACGTCGTCGTCATTAACCTCAGCAATCCTAAATCTGACGCCGAGCTAGGATAGCT 401
QY 343 TCATGAACCCAAATTCAGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
Db 402 TCATGAACCCAAATTCAGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461
QY 403 AGACAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462
Db 462 AGACAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 521
QY 463 CTCGCCCTTCCTCCCTCCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 522
Db 522 CTCGCCCTTCCTCCCTCCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 581
QY 523 CAATATGACCTCCTCTATGCTGTGGCCAACTAGGACGAGAGAGAGAGAGAGAGAG 582
Db 582 CAATATGACCTCCTCTATGCTGTGGCCAACTAGGACGAGAGAGAGAGAGAGAGAG 641
QY 583 TGCAGGAGCTGAGTCACCCCAAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 642
Db 642 TGCAGGAGCTGAGTCACCCCAAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 701
QY 643 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
Db 702 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 761
QY 703 GCCACCTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 762

```

```

Db 762 GCCACCTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 821
QY 763 ATCATGCTTTTCCCTCCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 822
Db 822 ATCATGCTTTTCCCTCCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 881
QY 823 GGTGTTTTCAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 882
Db 882 GGTGTTTTCAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 941
QY 883 GATATGTCCTCTCATATCATATACCCATTAGTATACCCATTATTTAGAA 934
Db 942 GATATGTCCTCTCATATCATATACCCATTAGTATACCCATTATTTAGAA 993

RESULT 9
AX365327
LOCUS 3184 bp DNA linear PAT 15-FEB-2002
DEFINITION Sequence 17 from Patent WO0204491.
ACCESSION AX365327
VERSION AX365327.1 GI:18697053
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE 1
AUTHORS Williams, S.R. and Rothermel, B.
TITLE Methods and compositions relating to muscle selective calcineurin
interacting protein (mcip)
JOURNAL Patent: WO 0204491-A 17 17-JAN-2002;
Board of Regents, The University of Texas System (US); Williams,
Sanders R. (US); Rothermel, Beverly (US)
FEATURES
source
Location/Qualifiers
1..3184
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
205..783
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23813.1"
/db_xref="GI:18697054"
/translation="MDCDVSTLVACVYDVEVFTNOEVEKFGGLFRTYDDCVTFOLFK
SFRVRINFSNPKSAARIEHETQFRGKILKLYFAOVOTPETDGLHLAPQPAK
QFLISPPSPVPVSWQPINDAITPVLYNDLYAVAKLGPGEKHELHAGTESTPSVVHVC
DSDIEEDPKTSPKPKIIQTRRPGLPSPVSN"
BASE COUNT 921 a 681 c 657 g 925 t
ORIGIN
Query Match 77.2%; Score 720.8; DB 6; Length 3184;
Best Local Similarity 95.9%; Pred. No. 1.4e-217;
Matches 740; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 163 CTCGTCGTCCTGTCGTCATGTCACAGTCAGTGTGTTGAAGGAGAGAGAGAGAA 222
Db 222 CACTCTGTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 281
QY 223 ATTTGAGGAGCTGTTTCGGAGCTTATGATGACTGTGTGACGTTCCAGCTATTTAAGAGTTT 282
Db 282 ATTTGAGGAGCTGTTTCGGAGCTTATGATGACTGTGTGACGTTCCAGCTATTTAAGAGTTT 341
QY 283 CAGACGTCGTCGTCATTAACCTCAGCAATCCTAAATCTGACGCCGAGCTAGGATAGCT 342
Db 342 CAGACGTCGTCGTCATTAACCTCAGCAATCCTAAATCTGACGCCGAGCTAGGATAGCT 401
QY 343 TCATGAACCCAAATTCAGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
Db 402 TCATGAACCCAAATTCAGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461
QY 403 AGACAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462

```

462 AGAGACAGATGGAGACAAACATGCACTTGGCTCCACCCAGCGCTGCCAAACAGTTTCTCAT 521  
463 CTCGCCCTCTCTCCACACCTGTTGGCTGGCAGCCCATCAACGATGCCAGCGAGTCTCT 522  
522 CTCGCCCTCTCTCCACACCTGTTAGCTGGCAGCCCATCAACGATGCCAGCGAGTCTCT 581  
523 CAACTATGACCTCTCTATGCTGGCCAAACTAGGACCGAGGAGAGATGAGCTCCA 582  
582 CAACTATGACCTCTCTATGCTGGCCAAACTAGGACCGAGGAGAGATGAGCTCCA 641  
583 TGCAGGAGTGGTCCACCCCAAGTGTCTGCTGCGACGTGCGACAGTGCACATAGAGGA 642  
642 TGCAGGAGTGGTCCACCCCAAGTGTCTGCTGCGACGTGCGACAGTGCACATAGAGGA 701  
643 AGAAGAGGACCCAAAGACTTCCCAAGCCAAATATCATCCAACTCGGCGTCTCGCCT 702  
702 AGAAGAGGACCCAAAGACTTCCCAAGCCAAATATCATCCAACTCGGCGTCTCGCCT 761  
703 GCCACCTCGGTCCTCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762  
762 GCCACCTCGGTCCTCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 821  
763 ATCATGCTTTTCCCGCTGTTTCTCAAAATAATGCTTTAAATTCCTGGGTGTTT 822  
822 ATCATGCTTTTCCCGCTGTTTCTCAAAATAATGCTTTAAATTCCTGGGTGTTT 881  
823 GGTGTTGAGATCTCTCTGTTATCAAGCTCTCGGACAAAGGCTAGGAAAGGT 882  
882 GGTGTTGAGATCTCTCTGTTATCAAGCTCTCGGACAAAGGCTAGGAAAGGT 941  
883 GATATGCTCTCTCATCATATCATATCATATCATATCATATCATATCATATCATATCAT 934  
942 GATATGCTCTCTCATCATATCATATCATATCATATCATATCATATCATATCATATCAT 993

RESULT 10  
AX420436 3184 bp DNA linear PAT 18-JUN-2002  
LOCUS AX420436 Sequence 12 from Patent WO0204513.  
DEFINITION AX420436  
ACCESSION AX420436.1 GI:21524589  
VERSION  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Loring, J.F., Tingley, D.W., Edwards, C.M. and Streeter, D.G.  
Down syndrome critical region 1-like 1 proteins  
Patent: WO 0204513-A 12 17-JAN-2002;  
Incyte Genomics, Inc. (US)  
Location/Qualifiers  
1..3184  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: g1435039"  
BASE COUNT 921 a 681 c 657 g 925 t  
ORIGIN

Query Match 77.2%; Score 720.8; DB 6; Length 3184;  
Best Local Similarity 95.9%; Pred. No. 1.4e-217;  
Matches 740; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 163 CTCGCTGTTGGTGGTCAATGTTCCAGCTGCTGTTGAAGGAGAGAGAGAGAGAGAGAGAG 222  
222 CACTCTGGTGGCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 281  
223 ATTTGAGGAGCTGTTTCCGAGCTTATGATGCTGTTGACGTTCCAGCTATTTAAGAGTTT 282  
282 ATTTGAGGAGCTGTTTCCGAGCTTATGATGCTGTTGACGTTCCAGCTATTTAAGAGTTT 341  
283 CAGACGCTGCTCGGTATTAACATTCAGCAATCTCTAAATCTCGACCGCGAGCTAGGATAGCT 342

342 CAGAGCTGCTCGTATAAATTCAGCAATTCCTAAATCTGCAGCCGAGCTAGGATAGAGCT 401  
343 TCATGAAACCAATTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402  
402 TCATGAAACCAATTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461  
403 AGACACATGAGAGACAAATTCGCTTGGCTCCACCCAGCGCTGCCAAACAGTTTCTCAT 462  
462 AGACACATGAGAGACAAATTCGCTTGGCTCCACCCAGCGCTGCCAAACAGTTTCTCAT 521  
463 CTCGCCCTCTCTCCACACCTGTTGGCTGGCAGCCCATCAACGATGCCAGCGAGTCTCT 522  
522 CTCGCCCTCTCTCCACACCTGTTAGCTGGCAGCCCATCAACGATGCCAGCGAGTCTCT 581  
523 CAACTATGACCTCTCTATGCTGCGCCAAACTAGGACCGAGGAGAGATGAGCTCCA 582  
582 CAACTATGACCTCTCTATGCTGCGCCAAACTAGGACCGAGGAGAGATGAGCTCCA 641  
583 TGCAGGAGTGGTCCACCCCAAGTGTCTGCTGCGACGTGCGACAGTGCACATAGAGGA 642  
642 TGCAGGAGTGGTCCACCCCAAGTGTCTGCTGCGACGTGCGACAGTGCACATAGAGGA 701  
643 AGAAGAGGACCCAAAGACTTCCCAAGCCAAATATCATCCAACTCGGCGTCTCGCCT 702  
702 AGAAGAGGACCCAAAGACTTCCCAAGCCAAATATCATCCAACTCGGCGTCTCGCCT 761  
703 GCCACCTCGGTCCTCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762  
762 GCCACCTCGGTCCTCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 821  
763 ATCATGCTTTTCCCGCTGTTTCTCAAAATAATGCTTTAAATTCCTGGGTGTTT 822  
822 ATCATGCTTTTCCCGCTGTTTCTCAAAATAATGCTTTAAATTCCTGGGTGTTT 881  
823 GGTGTTGAGATCTCTCTGTTATCAAGCTCTCGGACAAAGGCTAGGAAAGGT 882  
882 GGTGTTGAGATCTCTCTGTTATCAAGCTCTCGGACAAAGGCTAGGAAAGGT 941  
883 GATATGCTCTCTCATCATATCATATCATATCATATCATATCATATCATATCATATCAT 934  
942 GATATGCTCTCTCATCATATCATATCATATCATATCATATCATATCATATCATATCAT 993

RESULT 11  
HUMZAKI4  
LOCUS HUMZAKI4 3184 bp mRNA linear PRI 06-FEB-1999  
DEFINITION HUMZAKI4  
ACCESSION D83407  
VERSION D83407.1 GI:1435039  
KEYWORDS ZAKI-4; thyroid hormone responsive.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 3184)  
Miyazaki, T., Kanou, Y., Murata, Y., Ohmori, S., Niwa, T., Maeda, K.,  
Yamamura, H. and Seo, H.  
Molecular cloning of a novel thyroid hormone-responsive gene,  
ZAKI-4, in human skin fibroblasts  
J. Biol. Chem. 271 (24), 14567-14571 (1996)  
862924  
2 (bases 1 to 3184)  
Miyazaki, T.  
Unpublished  
3 (bases 1 to 3184)  
Miyazaki, T.  
Direct Submission  
Submitted (06-FEB-1996) Takashi Miyazaki, Research Institute of  
Environmental Medicine, Nagoya Univ., Department of Endocrinology  
and Metabolism, Furo-chou, Chikusa-ku, Nagoya, Aichi 464-01, Japan  
(E-mail:tama@iem.nagoya-u.ac.jp, Tel:052-789-3867,

Query Match	Score	DB 9;	Length
Best Local Similarity	95.9%;	Pred. No. 1.4e-217;	
Matches	740;	Conservative	0; Mismatches 32; Indels 0; Gaps 0;
Qy	163	CTCGTGTGTTGGTCCAAATGTTCCACAGTCAGTCTTTGAAGGAGAAAGCAGCAAGAAA	222
Db	222	CACTCTGGTTCCTGTGTGGGATGCGAGGTCTTTACCAATCAGGAGTTTAAGGAAA	281
Qy	223	ATTTCAGGAGCTGTTTCGGACTTATGATGACTGTGTCAAGTTCACAGTATTTTAAGAGTTT	282
Db	282	ATTTGGGGGACTGTTTCGGACTTATGATGACTGTGTCAAGTTCACAGTATTTTAAGAGTTT	341
Qy	283	CAGAGGTGTCGGTATAAAGTTCAGCAATCTTAATCTGCAGCCCGAGTCAGGATAGAGCT	342
Db	342	CAGAGGTGTCGGTATAAAGTTCAGCAATCTTAATCTGCAGCCCGAGTCAGGATAGAGCT	401
Qy	343	TCATGAACCAATTCAGAGGGGAAAAATTAAGTCTCTACTTTGCACAGGTTCAGACTCC	402
Db	402	TCATGAACCAATTCAGAGGGGAAAAATTAAGTCTCTACTTTGCACAGGTTCAGACTCC	461
Qy	403	AGACACAGATGGAGACAAACTGCACCTTGGCTCCACCCAGGCTGCACCAACAGTTTCAT	462
Db	462	AGACACAGATGGAGACAAACTGCACCTTGGCTCCACCCAGGCTGCACCAACAGTTTCAT	521
Qy	463	CTCGCCCTTCCTCCCACTGTTGGCTGGCAGCCCATCAACCATGCCAGGCCAGTCCT	522
Db	522	CTCGCCCTTCCTCCCACTGTTGGCTGGCAGCCCATCAACCATGCCAGGCCAGTCCT	581
Qy	523	CAACTATGACCTCTCTATGCTGTGGCCAAACTAGCAGGAGAGAAATGATGAGTCCCA	582
Db	582	CAACTATGACCTCTCTATGCTGTGGCCAAACTAGCAGGAGAGAAATGATGAGTCCCA	641
Qy	583	TGCAGGGACTGAGTCCACCCCAAGTGTGCTGTCAGGTGTGCAGCAGTGACATAGAGGA	642
Db	642	TGCAGGGACTGAGTCCACCCCAAGTGTGCTGTCAGGTGTGCAGCAGTGACATAGAGGA	701
Qy	643	AGAAGAGACCAAGACTTCCCAAGCCAAAAATCATCCAACTCGCGCTCCGCTCT	702
Db	702	AGAAGAGACCAAGACTTCCCAAGCCAAAAATCATCCAACTCGCGCTCCGCTCT	761
Qy	703	GCCACCTTCGTTCCAACTGAGTCGCCGTGCTCTTCGATAAATAGCGCTCTCCTCTTT	762
Db	762	GCCACCTTCGTTCCAACTGAGTCGCCGTGCTCTTCGATAAATAGCGCTCTCCTCTTT	821

## FEATURES

RESULT 13  
AF459023 953 bp mRNA linear  
LOCUS AF459023 complete cds.  
beta mRNA.

Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Muridae; Murinae;

REFERENCE  
AUTHORS  
1 (bases 1 to 953)  
Miyazaki, T., Cao, X., Kambe, F., Ohmori, S. and Seo, H.  
Cloning of rat ZAKI-4 beta CDNA

JOURNAL  
Unpublished  
2 (bases 1 to 953)  
REFERENCE  
Miyazaki, T., Cao, X., Kambe, F., Ohmori, S. and Seo, H.  
AUTHORS  
Direct Submission  
TITLE  
Submitted (13-DEC-2001) Endocrinology and Metabolis  
JOURNAL

**AUTHORS** Miyazaki, T., Cao, X., Ramberg, J.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (13-DEC-2001) Endocrinology and Metabolism, Research

## FEATURES

DS

BASE COUNT  
ORIGIN

Query Mat  
Best Loca  
Matches

Ov

nb

Oy

Dbb

QY

Db

QY

99

59

3

3

3

2

Qy

Dò

oy

db

QY

1

Db	739	TTCCCCAAGCCAAAAATCATCCAGACCCGGCTCTGGCCTGCCACCTCCGTGTC	798
Qy	721	CTGAGCTGCCGTCTTCGATATAGCCGTCTCTTTATCATGCTTTTCCCT	780
Db	799	CTGA-----GCAGTCCATGATAACATGCATCTCTC-TTATCATGCTTTTCCCT	850
Qy	781	GTGCTTTGTCAAAAAATTCCTTTAAATTCCTGGTGTGTGGTGTTCGAGATCCTT	840
Db	851	TGTTGCGTCCAAAAATTAATGCTTGAAGTCCTCGGTGTTGGTGTTCGAGAATCCTT	910
Qy	841	CCTGTGTATCAAGCCCTCGGACAAAGGGCTAGAAAAGGTG	883
Db	911	CCTTGTAATTAGCCCTCTCGGACAAAGGGCTAGAAAAGGTG	953

[illegible]

SOURCE	ORGANISM
	Mus musculus (house mouse)
	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1. (bases 1 to 3080)

Submitted (21-FEB-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590  
USA

REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Issue Procurement: The Cepko Laboratory  
CDNA Library preparation, 1988

CDNA Library Arrayed by: The I.M.A.G.E. Consortium  
Preparation: Life Technologies, Inc.

DNA sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)

Gaithersburg, Maryland;

web site: <http://www.nisc.nih.gov/>  
Contact: nisc\_macc@nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S. V., Beckstrom, R. L., and Beckstrom, R. L. 1991. *Journal of the American Academy of Dermatology* 25: 1111-1115.

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, C., Dietrich, N.I., Gorman, J., Hargrave, S.M., Benjamin, B., and O'Brien, S.M.

Hansen, N., Ho, S.-L., Karlins, E., and Karlins, P., 1981, *Journal of Polymer Science, Part A: Polymer Chemistry*, **19**, 1001-1010.

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McClellan, J., McEwen, L., Kwong, P., Laric, P., Legaspi, R., McDowell, J., and Mastrian, S.D. (1994) The effects of chronic stress on the immune system. *Stress* 17: 1-10.

McBowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
Tsurgeon, C., Vogt, J., Walker, J., Thomas, P.J., Touchman, J.W.

Young, A., Zhang, L.-H. and Green, E.D.  
Young, A., Zhang, L.-H. and Green, E.D.  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MCC clone 3: 1

through the I.M.A.G.E. Consortium/LLNL at: <http://www.llnl.gov/imag>. Information can be found

Series: IRAX Plate: 100 Row: n Column: 10  
This clone was selected from

was selected for full length sequencing because it passed the following selection criteria: matched the

Location/Qualifiers	source
1 3090	

```
1: 2000
/organism="Mus musculus"
```

```

/mol_type="mRNA"

```

```

/uu_xref="taxon:10090"
/clone="IMAGE:6491318"

```

```

/tissue_type="Eye, retina, mouse strain C57BL/6"

```

```
/clone_lib="NIH_MGC_94"  
/lab_host="DH10B"
```

```

E COUNT
      /note="Vector: pCMV-SPORT6"
      /name="Durob"

```

COGNOME	NUMERO	SESSO	ANNO DI NASCITA	ANNO DI MORTE
COGNOME	882 a	677 c	678 g	843 t
GIN				

Query Match

1000

\_\_\_\_\_

Best Local Similarity 79.9%; Pred. No. 6.6e-152; Matches 618; Conservative 0; Mismatches 154; Indels 1; Gaps 1;	
QY	163 CTCGTGTTGTTGGCGTCAATGTTTCACCACTCAGTGTGTTGAAGAGAGAGACGACGAA 222
Db	91 CACTCTGTCGCTGTGTGGTGGATGTGGAGGTCTTTACCAATCAGGAGTTAAGGAAA 150
QY	223 ATTTGAGGACTGTTTCGGACTTATGATGACTGTGTGACGTGTCCAGCTATTATAGAGTTT 282
Db	151 ATTTCGAGGACTGTTTCGGGACCTATGATGAATGTGTGACGTGTCCAGCTGTTTAAAGAGTTT 210
QY	283 CAGACGTGTCCGTATAAATCTCAGCAATCTTAATCTGCAGCCGAGCTAGGATAGAGCT 342
Db	211 CCGACGGTTCGAATAAATTCAGCCATCCCAATCTGCAGCCGCTGCCCGATAGAGCT 270
QY	343 TCATGAACCAATTCAGAGGGAATAAATTAAGCTCTACTTTGCACAGGTTTCAGACTCC 402
Db	271 TCATGAGACTCAGTTCAGAGGGAAGAAGCTAAAACCTACTTCCCGCAGGTCGACACCC 330
QY	403 AGACAGAGATGGAGACAACTGCACCTGGCTCCACCCAGCCCTGCCAAACAGTTTCTCAT 462
Db	331 AGACAGAGATGGAGACAACTGCTTTGGCACCTCCACAGCCTGCCAAACAGTTCTCAT 390
QY	463 CTCGCCCTCTCTCCTCCCACTTGTGGCTGGCAGCCCATCAACGATGCCACGCCAGTCT 522
Db	391 CTCACCCCTCTCATCTCTCTCTCTGTTGGCTGGAAGCTATCAGGATGCCACACCACTCT 450
QY	523 CAACATGACCTCTCTATGCTGGCCAACTAGGACCGAGAGAAATATGAGTCCA 582
Db	451 CAACATGACCTCTTATGCTGGCCAACTAGGACCGAGAGAAATATGAGTCCA 510
QY	583 TGCAGGACTGAGTCCACCCCAAGTGTGCTGTGCACGTGTGCACAGTGCACATAGAGA 642
Db	511 CGCTGGAATGAGTCTACACGAGGCTGTGTGTCATGTGTGACCGACATAGGAGA 570
QY	643 AGAAGGAGCCCAAGACTTCCCAAGCCAAAATCATCCAACTCGGCGCTCTGGCT 702
Db	571 GGAGGAGGCCCAAGACTTCCCAAGCCAAAATCATCCAGACCCGCGCTCCGCGCT 630
QY	703 GCCACCTCTCGTGTCCAACGTAGC - TGCCCTGTCTCTTCGTGATATAGCCGTCTCCTCT 761
Db	631 GCCACCTCTCGTGTCCAACGTAGC - TGCCCTGTCTCTTCGTGATATAGCCGTCTCCTCT 690
QY	762 TATCATGCTTTTTCCTCCCTCTGTGTTGTCAAAAAAATTCGCTTTAAATTCCTGGGTGT 821
Db	691 TCCTCGCCCTCCCTCTGTGTTGTCCAAACAATTAATTCGCTTTAAGTCCCTGGGTGT 750
QY	822 TGGTGTGTTGAGATTCCTCTCTGTTTATCAAGCTCTCGGACAAAAGGCTAGGAAAAG 881
Db	751 TGGTGTGTTGGAATCTCTCTGTTGTAATTAAGCTCTCGGACAAAAGGCTAGGAAAAG 810
QY	882 TGATATGCTCTCTGATCATATACCCCAATTAAGTATACCCATTAATTTAGAA 934
Db	811 TGATATGCTCTCCAATCACTCTACCCCAATTAAGTATACCCATTAATTTAGAA 863

RESULT 15					
BC049096					
LOCUS	BC049096	3263 bp	mRNA	linear	ROD 21-MAR-2003
DEFINITION	Mus musculus, Down syndrome critical region gene 1-like 1, clone				
ACCESSION	MGC:61350 IMAGE:6416901, mRNA, complete cds.				
	BC049096				

BC043030	GI:29145001
VERSION	BC043036.1
KEYWORDS	MGC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 3263)  
Strausberg, R.  
Direct Submission  
Submitted (19-MAR-2003), National Center for Biotechnology Information

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Query Match	Score	DB 10;	Length	3263;
Best Local Similarity	79.9%;	Pred. No. 6.6e-152;	Indels	1; Gaps
Matches	618;	Conservative	0; Mismatches	154;
163	CTCGTTGTTTGGCTGCAATGTTTACCAGTCAGTGTGTTGAAGGAGAAAGAGAGCAGCAAGGAAA	222	1	1
1	1	1	1	1
293	CACCTCTGGTCGCTGTGTGGTGTGATGTCGGAGCTCTTTACCAATCAGGAGGTTTAAGGAAA	352	1	1
1	1	1	1	1
293	ATTTGAGGGAGCTGTTTCGGACTTATGATGACTGTGTGACGTTCCAGCTTATTAAGAGTTT	282	1	1
1	1	1	1	1
353	ATTCGAGGAGCTGTTTCGGACCTATGATGAATGTGTGACGTTCCAGCTGTTTAAGAGTTT	412	1	1
1	1	1	1	1
283	CAGACGTTCCGTTATAAATTCAGCAATCTCTAAATCTGCAGCCCGAGCTAGGATAGAGCT	342	1	1
1	1	1	1	1
413	CCGAGCGGTTTCGAATAAATTTACGCCATCCCAATCTGCAGCCCGTGC CGGATAGAGCT	472	1	1
1	1	1	1	1
343	TCATGAAACCCAAATTCAGAGGGAAAAAATTTAAAGCTCTACTTTGCACAGGTTTCAGACTCC	402	1	1
1	1	1	1	1
473	TCATGAGACTCAGTTTCAGAGGGAAGAAGCTAAAACCTCTACTTCGCCCGAGTCCAGACCCC	532	1	1
1	1	1	1	1
403	AGAGACAGATGAGACAACCTGCATTTGGGTCCACCCGCCCTGCCAAACAGATTTCTCAT	462	1	1
1	1	1	1	1
533	AGAGACAGATGAGACAACCTGCATTTGGCACCCTCCACAGCCCTGCCAAACAGATTTCTCAT	592	1	1
1	1	1	1	1
463	CTCGCCCTTCTCCTCCCAACCTGTTGGTGGCAGCCCATCAACGATGCCACGCCAGTCTCT	522	1	1
1	1	1	1	1
593	CTCACCCCTTTCATCTCCTCTGTTGGTGGGAAGCCCTATCAGCGATGCCACACAGTCTCT	652	1	1
1	1	1	1	1
523	CAACTATGACCTCCCTATGCTGTGGCCAAACTAGGACACGAGAGAGAGTATGAGCTCCA	582	1	1
1	1	1	1	1

Search completed: August 24, 2003, 11:27:03  
Job time : 4686 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 18:54:42 ; Search time 29 Seconds  
(without alignments)  
341.405 Million cell updates/sec

Title: US-10-030-613-1

Perfect score: 1254

Sequence: 1 MRSFGQGHVPEGGFLC.....SPKLIQTRRGLPPSVSN 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1254	100.0	255	4	US-09-614-474-2
2	913	72.8	192	4	US-09-614-474-10
3	729.5	58.2	241	4	US-09-614-474-11
4	581	46.3	171	2	US-08-665-040-2
5	98	7.8	454	3	US-08-348-518C-4
6	98	7.8	454	3	US-08-476-509B-4
7	94	7.5	905	3	US-08-574-959A-9
8	94	7.5	905	3	US-09-357-014-9
9	94	7.5	1135	3	US-08-574-959A-7
10	94	7.5	1135	3	US-09-357-014-7
11	90	7.2	655	3	US-09-347-833-4
12	83.5	6.7	659	4	US-09-562-737-12
13	83	6.6	554	4	US-09-328-352-7796
14	82	6.5	243	1	US-07-960-112B-2
15	82	6.5	243	1	US-08-301-316B-2
16	82	6.5	243	2	US-08-473-399B-2
17	82	6.5	243	3	US-08-853-831-2
18	82	6.5	243	4	US-09-510-885-2
19	82	6.5	243	5	PCR-US93-09774-2
20	81.5	6.5	906	4	US-09-651-656-13
21	81.5	6.5	906	4	US-09-650-855-13
22	80.5	6.4	415	4	US-09-252-991A-31348
23	80.5	6.4	830	4	US-09-252-991A-22004
24	80	6.4	428	3	US-09-347-833-6
25	79.5	6.3	424	4	US-09-333-593A-8
26	79.5	6.3	1706	2	US-08-459-568-2
27	79.5	6.3	1706	2	US-08-399-411-2

28 79.5 6.3 1706 3 US-08-516-859A-2 Sequence 2, Appli  
29 79.5 6.3 1706 4 US-09-586-472-2 Sequence 2, Appli  
30 79.5 6.3 1706 4 US-09-528-706-2 Sequence 2, Appli  
31 79 6.3 434 2 US-08-815-718-3 Sequence 3, Appli  
32 79 6.3 529 4 US-09-252-991A-18830 Sequence 18830, A  
33 79 6.3 529 2 US-08-884-072-6 Sequence 6, Appli  
34 79 6.3 559 4 US-09-212-168-6 Sequence 6, Appli  
35 78.5 6.3 580 4 US-09-252-991A-29484 Sequence 29484, A  
36 77.5 6.2 350 2 US-08-960-022-8 Sequence 8, Appli  
37 77.5 6.2 782 4 US-09-252-991A-30464 Sequence 30464, A  
38 77 6.1 681 6 5194595-19 Patent No. 5194595  
39 76.5 6.1 330 3 US-09-145-391-2 Sequence 2, Appli  
40 76 6.1 566 4 US-09-252-991A-16790 Sequence 16790, A  
41 76 6.1 714 4 US-09-347-878-16 Sequence 16, Appli  
42 75 6.0 930 4 US-09-254-594-6 Sequence 6, Appli  
43 75 6.0 2441 1 US-08-194-468-2 Sequence 2, Appli  
44 75 6.0 2441 3 US-08-961-739-2 Sequence 2, Appli  
45 75 6.0 2441 4 US-09-514-247A-8 Sequence 8, Appli

#### ALIGNMENTS

RESULT 1  
US-09-614-474-2  
; Sequence 2, Application US/09614474  
; Patent No. 6524819

GENERAL INFORMATION:

APPLICANT: Jor Ing, Jeanne F.

APPLICANT: Tingley, Debora W.

APPLICANT: Edwards, Carla M.

APPLICANT: Streeter, David G.

TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS

FILE REFERENCE: PG-0013 US

CURRENT APPLICATION NUMBER: US/09/614,474

CURRENT FILING DATE: 2000-07-11

NUMBER OF SEQ ID NOS: 13

SOFTWARE: PERL Program

SEQ ID NO 2

LENGTH: 255

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No. 6524819 247500.5

NAME/KEY: unsure

LOCATION: 11

OTHER INFORMATION: unknown or other

US-09-614-474-2

Query Match

Best Local Similarity 100.0%; Score 1254; DB 4; Length 255;

Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSFGQGHVPEGGFLCICIDRDWAVTRCFEAFAQIITDNDLNSLFACNVHOSVF 60

Db 22 MRSFGQGHVPEGGFLCICIDRDWAVTRCFEAFAQIITDNDLNSLFACNVHOSVF 81

QY 61 EGESKEKEFLRTYDDCVTFQLFKSFRVRIRNFSNPKSAARARIELHETQFRGKKLKL 120

Db 82 EGESKEKEFLRTYDDCVTFQLFKSFRVRIRNFSNPKSAARARIELHETQFRGKKLKL 141

QY 121 YFAQVQTPETDGDGLHAPPQAKOFLISPPSPVGVQPIINDATPVNLVLLYAVAKLG 180

Db 142 YFAQVQTPETDGDGLHAPPQAKOFLISPPSPVGVQPIINDATPVNLVLLYAVAKLG 201

QY 181 PGEKYLHAGTSPSVVHVCDSDIEEDPKTSKPKLIQTRRGLPPSVSN 234

Db 202 PGEKYLHAGTSPSVVHVCDSDIEEDPKTSKPKLIQTRRGLPPSVSN 255

RESULT 2

US-09-614-474-10

[illegible][illegible]

```

128 PETDGDKLLHAPPQAKOFLISPPSSPPVQWQPIN DATPVNLYDLLYAVAKLGGPKYEL 187
      :      ||||| ||||| ||||| :      ||||| ||||| ||||| ||||| |||||
65  IGSS-----HLAPNPDKOFLISPPASPVYQVEDATPVNLYDLYAKSLGKGPKYEL 120
      :      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
188 HACTESTSPVVHVHCDSDIEEDPKTS---PKPKIITOTRRGLPP 230
      || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 HAATDTTSPVVHVHVCSDQKEEEMEMRRPKIITOTRRPYTP 167
      || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 5  
 US-08-348-518C-4  
 ; Sequence 4, Application US/08348518C  
 ; Patent No. 6022740  
 ; GENERAL INFORMATION:  
 APPLICANT: SUDOL, MARIUS  
 APPLICANT: PEER, BORK  
 APPLICANT: HENRY, CHEN  
 TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A  
 SIGNALLING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE  
 PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC  
 USES THEREOF  
 NUMBER OF SEQUENCES: 26

ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/348.518C  
FILING DATE: 03

DATE: 01-DEC-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-101  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 454 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
08-348-518C-4

```

very Match          7.8%; Score 98; DB 3; Length 454;
st Local Similarity 31.1%; Pred. No. 0.012;
atches 33; Conservative 11; Mismatches 30; Indels 32; Gaps 5;

139 PPQPAKQFLISPPSPVGVQPIINDATPVNLDDLYAVAKLGPGEKYELHAGTESPPSV- 197
||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
8 PPQAPAQGGQPPSGPPGQGPPS-----GPGQ--PAPAATQAAPQAP 48
||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

198 -----VVHV-CDSIDIEEE-----DPKTSPEKKIITOTRRGLPPS 231
|||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
49 PAGHOIVHVRGDSETDLEALFNAMNPKTANYPQTVPMLRKLPPS 94

```

T 6  
i-476-509B-4  
uence 4, Application US/08476509B  
NERAL INFORMATION:  
APPLICANT: SUDOL, MARIUS

APPLICANT: PEER, BORK  
 APPLICANT: HENRY, CHEN  
 TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A  
 TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE  
 TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC  
 TITLE OF INVENTION: THEREOF  
 NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Klauber & Jackson  
 STREET: 411 Hackensack Avenue  
 CITY: Hackensack  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07601  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/476.509B  
 FILING DATE: 01-DEC-1994  
 CLASSIFICATION:

TORREY/AGENT INFORMATION:  
 NAME: Jackson Esq., David A.  
 REGISTRATION NUMBER: 26,742  
 REFERENCE/DOCKET NUMBER: 600-1-101 CIP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201 487-5800  
 TELEFAX: 201 343-1684  
 TELEX: 133521  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 454 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 76-509B-4

```

Match
Local Similarity 7.8%; Score 98; DB 3; Length 454;
Gaps 33; Conservative 11; Mismatches 30; Indels 32; Gaps 5;

139 PPQAKOFLIIPSSPPVGMQINDATPVLYDLLVAVAKLCPGKTELIHAGTSTFSV- 197
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8 PPQAPQGGGGPPQSPQGGCGPTs-----GPGQ--PAPAAQAQAQAP 48

198 -----VVHV-CDSDIEEEE-----DPKTSPPKIIQTRRGLPPS 231
|||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
49 PAGHQIVIVHVRGDSQTDLEALFNVAVMKPTANYPQVTPVPMRLRLKLPDS 94

```

4-959A-9  
 Case 9, Application US/08574959A  
 No. 5962224  
 AL INFORMATION:  
 APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi  
 ADDRESS OF INVENTION: Jack L. Strominger  
 DATE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES  
 DATE OF INVENTION: AND USES THEREFOR  
 NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 60 State Street, Suite 510  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 TELEPHONE: 02109-1875  
 OTHER READABLE FORM:  
 DISC TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 DRAWING SYSTEM: PC-DOS/MS-DOS

RESULT 8  
 US-09-357-014-9  
 ; Sequence 9, Application US/09357014  
 ; Patent No. 6291645  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadlamudi  
 ; and Jack L. Strominger  
 ; TITLE OF INVENTION: PE2 POLYPEPTIDES, RELATED POLYPEPTIDES  
 ; AND USES THEREFOR  
 ; NUMBER OF SEQUENCES: 22  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 State Street, Suite 510  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109-1875  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION NUMBER: US/09/357,014  
 ; FILING DATE: 19-Jul-1999  
 ; PRIOR APPLICATION NUMBER: 08/574,959  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mandragouras, Amy E.  
 ; REGISTRATION NUMBER: 36,207  
 ; REFERENCE/DOCKET NUMBER: DFN-008  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617)227-7400  
 ; TELEFAX: (617)227-5941  
 ; INFORMATION FOR SEQ ID NO: 9:  
 ; SEQUENCE CHARACTERISTICS:

```

JS-08-574-959A-7
Query Match          7.5%; Score 94; DB 2; Length 1135;
Best Local Similarity 23.7%; Pred No. 0.14;
Matches 41; Conservative 20; Mismatches 80; Indels 32; Gaps 6;

61 EGEEKKFGFLRTYDDCVTQLFKSFRVINFNPKSAARIELHETQFGKKKL 120
   |||..||| : : : ||| :
QY -----EEEEEEEEEEVE----- 966

928 EEEEEEEFEFEFEEG-----EEEEEEEEEEEEEEVE----- 966
Db -----

121 YFAQVQPTDGDKLHLAPQPAKQFLL19PPSPVWGQPINATPVNLNYDL1AVAKLG 180
QY -----

```

Db 967 ---DLEFGTAGGEVSGAPPPTLPALPPSP-KVQPEPEP-----GILLEVEE--- 1016  
Qy 181 PGKYLHAGTSTPSVHVHCDSDIEE-EDPKTSPKPKIITRRPGLPPSV 232  
Db 1017 PGTEERGADTAPTLAEPALPSQGEVEREGESPAAGPPQQLVEEPEPSXPTL 1069

## RESULT 10

US-09-357-014-7  
; Sequence 7, Application US/09357014  
; Patent No. 6291645  
; GENERAL INFORMATION:  
; APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi  
; and Jack L. Strominger  
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES  
; AND USES THEREFOR  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/357,014  
; FILING DATE: 19-Jul-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/574,959  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: DFN-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1135 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Query Match 7.5%; Score 94; DB 3; Length 1135;  
Best Local Similarity 23.7%; Pred. No. 0.14;  
Matches 41; Conservative 20; Mismatches 80; Indels 32; Gaps 6;  
Qy 61 EGESKEKEGLRTYDDCVTFQLFKSFRRVFNPNKSAARARIELHETQFGKKLKL 120  
Db 928 EEEEEEEEEEFEEG-----ELEEEEEEEEEELEEEVE----- 966  
Qy 121 YFAQVQTPETDGLKHLAPPQAKQFLISPPSPVGVQWQIPNDATPVNLNLYAVAKLG 180  
Db 967 ---DLEFGTAGGEVSGAPPPTLPALPPSP-KVQPEPEP-----GILLEVEE--- 1016  
Qy 181 PGKYLHAGTSTPSVHVHCDSDIEE-EDPKTSPKPKIITRRPGLPPSV 232  
Db 1017 PGTEERGADTAPTLAEPALPSQGEVEREGESPAAGPPQQLVEEPEPSXPTL 1069

## RESULT 11

US-09-347-833-4  
; Sequence 4, Application US/09347833  
; Patent No. 6294658

; GENERAL INFORMATION:  
; APPLICANT: Famodu, Layo O.  
; APPLICANT: Odell, Joan T.  
; TITLE OF INVENTION: Factors Involved in Gene Expression  
; FILE REFERENCE: BB-1172  
; CURRENT APPLICATION NUMBER: US/09/347,833  
; CURRENT FILING DATE: 1999-07-02  
; EARLIER APPLICATION NUMBER: 60/092,415  
; EARLIER FILING DATE: July 10, 1998  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 4  
; LENGTH: 655  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-09-347-833-4

Query Match 7.2%; Score 90; DB 3; Length 655;  
Best Local Similarity 23.4%; Pred. No. 0.17;  
Matches 47; Conservative 25; Mismatches 89; Indels 40; Gaps 8;  
Qy 23 DRDNAVTRCFA-----EEAFOAITDNDLPSLFCACNVHOSVFEESEKKEFE 70  
Db 280 DKEWYVCRAQKKSEREMELKEFEKNIKEAADKNOGTN-LYLKLNLDSDIDDEKLKEIFA 338  
Qy 71 GLFRTYDDCVTFQLFKSFRRV-----INFSNPKSAARARIELHETQFGKKLKYAQV 125  
Db 339 D-FGTITSC---KVRDLNGVSGSGFVAFKSAEDASRALVAMNGKMTGSK--PLYVALA 392  
Qy 126 QTPETDGLKHL-----LAPPQAKQFLISPPSPVGVQWQIPNDATPVNLNLYAV 176  
Db 393 QRKEERRARLQAQFSQMRPMVMPSPVAPRPMYPPGVGVQQLFYQOPPP-----AF 445  
Qy 177 AKLGPGEKYLHAGTSTPSV 197  
Db 446 VNPQPGFGFQOHLIPGMRPSV 466

## RESULT 12

US-09-562-737-12  
; Sequence 12, Application US/09562737  
; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; APPLICANT: Gotthardt, Michael  
; TITLE OF INVENTION: LDL Receptor Signaling Pathways  
; FILE REFERENCE: UTSW0708  
; CURRENT APPLICATION NUMBER: US/09/562,737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 659  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-09-562-737-12

Query Match 6.7%; Score 83.5; DB 4; Length 659;  
Best Local Similarity 22.0%; Pred. No. 0.98;  
Matches 54; Conservative 29; Mismatches 82; Indels 81; Gaps 12;  
Qy 47 PNLFCACNVHOSVFEESEKKEFGLFRTYDDC-VTFQLEK--SFRRVRINFSNPKSAAR 103  
Db 30 PNGLRTHDISLEHFEDELDSE-----IIDCGISLQCLDTLSLRPRMGLLSAGSGN 82  
Qy 104 ARIELHETQFGKKLKYFAQVQTPETDGLKHLAPPQAKQFLISPPSPV----- 156  
Db 83 AGSRL---QAEMQQMDLIDAAAGRTPGAEDDESEDELAQAQRTGVGPPKAESVQDPAPRS 139  
Qy 157 -GWQP-----INDATPV-----LNYDLLYAVAKLGPEKYE----- 186



us-10-030-613-1.rai

```

QY      108 LHETQFRGKKLKYFAOYTPE-----TDGDKHLHLPQPAKQFLI 148
Db      43 LHE-----LYLDVTAPEDNEEASVQIFPDSVMLAVQEGDILLTFPPAGS---- 89
QY     149 SPPSPPPVGHPINDATPVLLYDLXAVAKLGKEKYLHA--CTESTPSVVHVCD--- 203
Db      90 --PEPHLSQP-----EQEQRALGPVSMNPLYVEIDLTC 125
QY     204 -----SDIEEDPKTSRKP-----LIQTRRGGLPPSV 232
Db     126 EAGFPSPDDEEGVPSEPEPEPEPEPARTRPKMAPAI 167

```

Search completed: August 22, 2003, 18:59:56  
Job time : 32 secs





Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	934	100.0	3159	4	US-09-614-474-1	Sequence 1, Appli
2	720.8	77.2	3184	4	US-09-614-474-12	Sequence 12, Appl
3	422	45.2	1021	4	US-09-614-474-3	Sequence 3, Appli
4	279.2	29.9	828	4	US-09-614-474-13	Sequence 13, Appl
5	257.6	27.6	615	4	US-09-614-474-8	Sequence 8, Appli
6	177.4	19.0	2174	2	US-08-665-040-1	Sequence 1, Appli
7	126.8	13.6	531	4	US-09-614-474-6	Sequence 6, Appli
8	43.6	4.7	297	4	US-09-614-474-9	Sequence 9, Appli
9	40.8	4.4	7218	1	US-08-232-463-14	Sequence 14, Appl
10	35	3.7	2254	2	US-08-635-066-1	Sequence 1, Appli
11	33	3.5	1011	4	US-09-328-352-3446	Sequence 3446, Ap
12	32.6	3.5	458	3	US-09-141-000-4	Sequence 4, Appli
13	31.6	3.4	15652	4	US-09-422-936-60	Sequence 60, Appl
14	31.4	3.4	2595	4	US-09-619-353-13	Sequence 13, Appl
15	31.4	3.4	98844	4	US-09-791-211-10	Sequence 10, Appl
16	31.2	3.3	1683	3	US-08-926-842B-36	Sequence 36, Appl
17	31.2	3.3	2735	4	US-09-551-974A-101	Sequence 101, App
18	31.2	3.3	3012	4	US-09-551-974A-94	Sequence 94, Appl
19	31.2	3.3	3134	2	US-08-533-669A-1	Sequence 1, Appli
20	31.2	3.3	3134	4	US-09-183-861-1	Sequence 1, Appli
21	31.2	3.3	3134	4	US-09-022-765-1	Sequence 1, Appli
22	31.2	3.3	3134	4	US-09-551-974A-1	Sequence 1, Appli
23	31.2	3.3	4233	4	US-09-551-974A-99	Sequence 99, Appl
24	31.2	3.3	4917	4	US-09-551-974A-100	Sequence 100, App
25	31.2	3.3	4929	4	US-09-551-974A-98	Sequence 98, Appl
26	31.2	3.3	10917	3	US-08-926-842B-11	Sequence 11, Appl
27	31	3.3	2157	4	US-08-973-005A-10	Sequence 10, Appl

301 CTTCCAGCAATCTAAATCTGCAGCCCGAGCTAGGATAGAGCTTCATGAAACCCAAATTCAG 360  
329 CTTCCAGCAATCTAAATCTGCAGCCCGAGCTAGGATAGAGCTTCATGAAACCCAAATTCAG 388  
361 AGGAAAAAATTAAGCTCTACTTTGACAGAGTTTCAGACTCCAGAGACAGATGGAGCAA 420  
389 AGGAAAAAATTAAGCTCTACTTTGACAGAGTTTCAGACTCCAGAGACAGATGGAGCAA 448  
421 ACTGCACTTGGCTCCACCCAGCTGCAAAACAGTTTCTCATCTCGCCCTTCTCTCCCT 480  
449 ACTGCACTTGGCTCCACCCAGCTGCAAAACAGTTTCTCATCTCGCCCTTCTCTCCCT 508  
481 ACCTGTTGGCTGGCAGCCCATCAACGATGCCAGCCAGTCTCAACTATGACCTTCTCTA 540  
509 ACCTGTTGGCTGGCAGCCCATCAACGATGCCAGCCAGTCTCAACTATGACCTTCTCTA 568  
541 TGTCTGGCCAACTAGGACAGGAGAGAGATGATGAGCTCCATGAGGAGTGGAGTCCAC 600  
569 TGTCTGGCCAACTAGGACAGGAGAGAGATGATGAGCTCCATGAGGAGTGGAGTCCAC 628  
601 CCAAGTGTGCTGCTGACAGTGTGGACAGTGCACATAGAGAGAGAGAGAGAGAGAGAG 660  
629 CCAAGTGTGCTGCTGACAGTGTGGACAGTGCACATAGAGAGAGAGAGAGAGAGAGAG 688  
661 TTCCCAAGCCAAATCATCAAACTCGGCGTCTGCGCTGCCAGCCCTCGGTGCCAA 720  
689 TTCCCAAGCCAAATCATCAAACTCGGCGTCTGCGCTGCCAGCCCTCGGTGCCAA 748  
721 CTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
749 CTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808  
781 GTTGTGTTGCAAAAAATTCCTTTAAATTCCTGGGTGTTGGTGTGTTGAGATTCCTT 840  
1009 GTTGTGTTGCAAAAAATTCCTTTAAATTCCTGGGTGTTGGTGTGTTGAGATTCCTT 1068  
841 CTTGTTTATCAAGCTCTCGACAAAGAGGCTAGGAAAAAGGTGATATGCTCTCTGATCAT 900  
1069 CTTGTTTATCAAGCTCTCGACAAAGAGGCTAGGAAAAAGGTGATATGCTCTCTGATCAT 1128  
901 ATCATACCAATTAAGTATACCCATTTAGAA 934  
1129 ATCATACCAATTAAGTATACCCATTTAGAA 1162

RESULT 2  
US-09-614-474-12  
; Sequence 12, Application US/09614474  
; Patent No. 6524819  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debora W.  
; APPLICANT: Edwards, Carla M.  
; APPLICANT: Streeter, David G.  
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS  
; FILE REFERENCE: PC-0013 US  
; CURRENT APPLICATION NUMBER: US/09/614,474  
; CURRENT FILING DATE: 2000-07-11  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PERL Program  
; SEQ ID NO 12  
; LENGTH: 3184  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Inocyte ID No. 6524819 g1435039  
US-09-614-474-12  
Query Match 77.2%; Score 720.8; DB 4; Length 3184;  
Best Local Similarity 95.9%; Pred. No. 1.2e-228;  
Matches 740; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

163 CTCGTTGTTCCGTCGCAATGTTCCACAGTCAGTCTTTTGAAGAGAGAGAGAGAGAGAG 222  
222 CACTCTGTTGCTGCTGTTGGATGTCGAGGCTCTTTACCAATCAGGAGGTTAAGAGAAA 281  
223 ATTTTCAGGAGGACTGTTTTCGAGCTTATGATGACTGTGTGACGTTCAGAGTATTTAGAGTTT 282  
282 ATTTTCGAGGACTGTTTTCGAGCTTATGATGACTGTGTGACGTTCAGAGTATTTAGAGTTT 341  
283 CAGAGCTGTGCTGATATAACTTTCAGCAATCTTAATCTGCAGCCCGAGCTAGGATAGAGCT 342  
342 CAGAGCTGTGCTGATATAACTTTCAGCAATCTTAATCTGCAGCCCGAGCTAGGATAGAGCT 401  
343 TCATGAAACCCCAATTCAGAGGAGAAAAATTAAGCTCTTACTTTTCACAGGTTTCAGAGTCC 402  
402 TCATGAAACCCCAATTCAGAGGAGAAAAATTAAGCTCTTACTTTTCACAGGTTTCAGAGTCC 461  
403 AGAGACAGATGAGAGACAACTGTCAGTCTGGCTCCACCCAGCTGCCAAACAGTTTCTCAT 462  
462 AGAGACAGATGAGAGACAACTGTCAGTCTGGCTCCACCCAGCTGCCAAACAGTTTCTCAT 521  
463 CTCGCCCCCTTCTCTCCACCTGTTGGCTGGCAGGCCATCAACGATGCCAGCCAGTCTCT 522  
522 CTCGCCCCCTTCTCTCCACCTGTTAGCTGGCAGGCCATCAACGATGCCAGCCAGTCTCT 581  
523 CAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCCAGGAGAGAGTATGAGCTCCA 582  
582 CAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCCAGGAGAGTATGAGCTCCA 641  
583 TGCAGGAGCTGAGTCCACCCCAAGTGTCTGCTGTCAGCTGTGCGACAGTGCATAGAGGA 642  
642 TGCAGGAGCTGAGTCCACCCCAAGTGTCTGCTGTCAGCTGTGCGACAGTGCATAGAGGA 701  
643 AGAAGAGGAGCCCAAGAGCTTCCCAAGCCAAATATCATCCAACTCGGCTCGGCTCT 702  
702 AGAAGAGGAGCCCAAGAGCTTCCCAAGCCAAATATCATCCAACTCGGCTCGGCTCT 761  
703 GCCACCTCGCTGTCACACTGAGCTGCTGCTCTCTCGATTAATAGCCGCTCTCTCTTT 762  
762 GCCACCTCGCTGTCACACTGAGCTGCTGCTCTCTCGATTAATAGCCGCTCTCTCTTT 821  
763 ATCATGCTTTTCCCTCTGTTTTCGAAAAAATTCGCTTTTAAATTCCTGGGTGTTT 822  
822 ATCATGCTTTTCCCTCTGTTTTCGAAAAAATTCGCTTTTAAATTCCTGGGTGTTT 881  
823 GGTGTTTTCGAGATTCCTCTCTCTTATCAAGCTCTCTCGACAAAGAGGCTAGGAAAAAGT 882  
882 GGTGTTTTCGAGATTCCTCTCTCTTATCAAGCTCTCTCGACAAAGAGGCTAGGAAAAAGT 941  
883 GATATGCTCTCTGATCATATCATACCCATTAAGTATACCCATTTATTAGAA 934  
942 GATATGCTCTCTGATCATATCATACCCATTAAGTATACCCATTTATTAGAA 993

RESULT 3  
US-09-614-474-3/c  
; Sequence 3, Application US/09614474  
; Patent No. 6524819  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debora W.  
; APPLICANT: Edwards, Carla M.  
; APPLICANT: Streeter, David G.  
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS  
; FILE REFERENCE: PC-0013 US  
; CURRENT APPLICATION NUMBER: US/09/614,474  
; CURRENT FILING DATE: 2000-07-11  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 1021  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus

```

; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: INCY
US-09-614-474-3

```

Query Match	45.28;	Score 422;	DB 4;	Length 1021;
Best Local Similarity	79.8%;	Pred. No. 9.1e-130;		
Matches 624;	Conservative 0;	Mismatches 140;	Indels 18;	Gaps 10;
QY	163	CTCGTTGTTGGCTGCAATGTTCCACGACTCAGTGTGTTGAAGGAGAGAGAGACGAGGAAA	222	
Db	823	CACTCTGGTGGCTGTGGTGATGTGGAGGTCTTTACCAATCAGGAGGTTTAAGGAAA	764	
QY	223	ATTGAGGAGACTGTTTCGGACTTATGATGACTGTGTGAGCTTCCAGCTATTTAAGAGTTT	282	
Db	763	ATTGAGGAGACTGTTTCGGAGCTATGACGAATGTGTGACGTTTTCAGCTGTTTAAGAGTTT	704	
QY	283	CAGAGCTG--TCCGTATAAATTCAGCAATCCTAAATCTGCAGCCG--AGCTAGATGAGA	339	
Db	703	CGAGCGGTTCGGAATAATTTAGCCACCCCAAGCTGCAGCCGCTGCCCGATAGGA	644	
QY	340	GCTTCA--TGAACCCCAATTCAGAGGGAAGAAAATTAAAGCTCTACTTTGCACAGGTTTCAG	397	
Db	643	GCTTCATGTGACCCAGCTTCATGAGGGAAGAAGCTGAAACTCTACTTTGCACAGGTTTCAG	584	
QY	398	ACTCCAGAGACAGATGG--AGACAAACTGCA--CTTGGCTCCACCCAGCCCTGCCAAGAGT	455	
Db	583	ACCCAGAGACAGATGTTAGCAAACTGCAATTTGGCACCCCAAGCTGCTGCCAAGAGT	524	
QY	456	TTCTCATCTCCGCCCTTCCTCCCACTGTTTGGCTGGCAGCCCACTCAACGATGCCACGC	515	
Db	523	TCCTCATCTACCCCTTCATCTCCCTCCGTTGGCTGGAAGCTTATCAGCGATGCCACAC	464	
QY	516	CAGTCCCTCAAC--TATGACCTCCCTCTATGCTGTGGCCCAACTAGCACAGAGAGAGATAT	574	
Db	463	CAGTCCCTCAAGGTACGACCTCCTTTATGCCGTGGCCCAACTAGCACAGAGAGAGATAT	404	
QY	575	GAGTCCCATCGAGGACTGAGTCCACCCCAAGTGTGCTGTGCAGCTGTGCGACAGTGCAC	634	
Db	403	GAGTGCATCGGGAACTGAGTCTACCCGAGCGTGTGCTGTGCAGCTGTGTGACAGGAC	344	
QY	635	ATAGAGGAAGAGGAGCACCAAGACTTCCCCAAAGCCAAAATATCATCAAACCTCGCGT	694	
Db	343	TTGGAGGAGGAGGAGGATCCAAAGACTTCCCCAAAGCCAAAATATCATCAAACCTCGCGT	284	
QY	695	CCTGGCTGCCACCTCCGCTGCCAAGTGCAGTGCCTTCCTTCGATTAATAGCCGTC	754	
Db	283	CCTGGCTGCCACCTCCGCTGCCAAGTGCAGTGCCTTCCTTCGATTAATAGCCGTC	231	
QY	755	TCCTCTTTATCATGCTTTTTCGCCCTGTGTTTGTCAAAAAAAATTCGCTTTAAAT--TCC	813	
Db	230	TCCTCT--TTATCATGCTTTTTCCTCCCTTGTTCGTCCTCAAAAATTAATTCCTTCAGTCACC	172	
QY	814	TGGGTGTTTGGTTTGGATTCCTTCCTTGTGTTA--TCAAGCCTCTCGGACAAAAGGCT	872	
Db	171	TGGGTATTTGGTTTGGAAATCTCTCCTTGTAACTTAAGCCTCTCGGACAAAAGGCT	112	
QY	873	AGGAAAGGTGATGTCCTCGTAFATATCATACCCATTAAAGTATATAACCCATTATTATAG	932	
Db	111	AGGAAAGGTGATGTCCTCGTCAACTCATACCCATTAAACAGAGTCCATTATTATAG	52	
QY	933	AA 934		
Db	51	AA 50		

RESULT 4  
US-09-614-474-13  
; Sequence 13, Application US/09614474  
; Patent No. 6524819  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debora W.

```

: APPLICANT: Edwards, Carla M.
: APPLICANT: Streeter, David G.
: TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS
: FILE REFERENCE: PC-0013 US
: CURRENT APPLICATION NUMBER: US/09/614,474
: CURRENT FILING DATE: 2000-07-11
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PERL Program
: SEQ ID NO 13
: LENGTH: 828
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID NO. 6524819 95017918
US-09-614-474-13

```

Query Match	29.9%;	Score 279.2;	DB 4;	Length 828;
Best Local Similarity	66.0%;	Pred. No. 2e-82;		
Matches 420;	Conservative 0;	Mismatches 213;	Indels 3;	Gaps 1;
QY	85	AGACAGGCACCTGGGCTGTCACCTCGTTGTTTGGATGCAAGAGAGCCCTTCAAGCAACTCACTGA	144	
Db	85	AGAGAGGAAGAAGAGATGATTTTGGTGAATGAAATGAAAGATGATTTGGATGAGATGATGGA	144	
QY	145	CTTCAATGACCTCCCCCAACTCGTTGTTTGGGTGCAATGTTCCACAGTCAGCTGTTTGAGG	204	
Db	145	TTTAAGTGATCGCTACCTCACCTCTTTGCTTGCAGCGTCCATGAAGCACTGTTTGAGGC	204	
QY	205	AGAAGAGCAAGGAAAAATTTGAGGACCTGTTTCGGACTTATGATGACTGTGTGACGTT	264	
Db	205	ACGAGAGCAGAAGAAAGATTTGAAGCACTCTTCACCATCTATGATGACCAGCTACTTT	264	
QY	265	CCAGCTATTTAAGAGTTTCAGACGTGTCCTGTAAATCTCAGCAATCCTAAATCTCGACG	324	
Db	265	TCAGCTGTTTAAAGCTTTAGAGAGTCAGATTAATTTTCAGCAAACTGTGAAGCGGCAGC	324	
QY	325	CCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAGAGGGGAAAAATTAAGGCTCTACTTT	384	
Db	325	AAGAGCCGGAATAGAATCCACGAAACAGACTTCAATGGCAGAACTAAAGCTATATTT	384	
QY	385	TGCACAGGTTTCAGACTCCAGAGACAGATGAGACAACTGCACCTTGGCTCCACCCGACGC	444	
Db	385	TGCACAGGTCAGATGTCCGGCGAAGTGGGGACAAGTCCTATCTCCTCCGCCCCGACGC	444	
QY	445	TGCCAAACAGTTTCTCATCTGCCCCCTTCTCTCCCACTGTGTGGCTGGCAGCCCATCA	504	
Db	445	TGTCAAGCAGTCTCTCATCTCCCTCCAGCCTCTCCCCAGTGGGTGGGAACAGACGCA	504	
QY	505	CGATGCCAGCCAGTCTCTCAACTATACCTCTCTATGCTGTGGCCAAACTAGGACCCAGG	564	
Db	505	AGATGCGATGCCGTGTTATAATATGATTTACTCTGTGCTGTTTCCAAATTTGGGACCAAGG	564	
QY	565	AGAGAAGTATGAGCTCCATGTCAGGGAAGTGTGAGTCCACCCCAAGTGTGCTGTGACGCTGTG	624	
Db	565	AGAGAATATGAACTTTCAGCGGGGAACAGAGTCGACCCACGCGTGGTGTTCATGTCTG	624	
QY	625	CGACAGTGCATAGAGGAAGAGAGGCCCAAGACTTCCCCAAAGCCAAATCATCCCA	684	
Db	625	TGAAGTGAACTAGAGAGGAAGAGAGCAAAAAA---ACCCCAACAGAAAAATTTGCCCA	681	
QY	685	AACTCGGCGTCTGGCGTGGCCACCCTCCGTGTCCAA	720	
Db	682	GCAAGGCGCCCGACCCCTCCGACCGCAGCGTTGAA	717	

RESULT 5  
US-09-614-474-8  
; Sequence 8, Application US/09614474  
; Patent No. 6524819  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debora W.

Tue Aug 26 09:11:05 2003

us-10-030-613-3.rni

APPLICANT: Edwards, Carla M.  
APPLICANT: Streeter, David G.  
TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS  
FILE REFERENCE: PC-0013 US  
CURRENT APPLICATION NUMBER: US/09/614,474  
CURRENT FILING DATE: 2000-07-11  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PERL Program  
SEQ ID NO 8  
LENGTH: 615  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No. 6524819 206751\_Rn.1  
US-09-614-474-8

Query Match 27.6%; Score 257.6; DB 4; Length 615;  
Best Local Similarity 68.9%; Pred. No. 2.5e-75;  
Matches 353; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

QY 119 GAAGAAGCCTTTCAAGCAATCACTGACTTCAATGACCTCCCACTGCTGTTGGTGGTGC 178  
DB 103 GAAGACGATTGGAGAGATGATGACCTCAGCGACCTGCCACCTCCCTCTTTGCTTGC 162  
QY 179 AATGTTCCAGTCAGTGTGTTGAAGGAGAGAGCAAGAAATTTGAGGAGCTGTTT 238  
DB 163 AGTGTCATGAAGCGGTGTTGAGGTTCCAAAGACAGAGAGGAGGTTCCAGAGCCCTGTT 222  
QY 239 CGGACTTATGATGACTGTGTGAGTTCAGCTATTTAAAGAGTTTCAGAGCTGTCCGTATA 298  
DB 223 ACCCTCTATGATGACAGGTCACATTCAGCTGTTCAAGAGCTTTCGACAGTGGAGATC 282  
QY 299 AACTTCAGCAATCTTAATCTGACGCCGAGCTAGGATAGAGCTTCATGAACCCCAATTC 358  
DB 283 AACTTCAGCAAGCTGAGGCTGCGCGGAGAGCAGGATCGAGCTCCACGAGAGTGATTC 342  
QY 359 AGAGGAAAAAATAAGCTCTACTTTGCACAGGTTTCAGACTCCAGAGACAGATGGAGAC 418  
DB 343 CATGCCCGGAGCTGAAGCTTACTTCGCACAGGTGCAGGTGTCGGGGAGGCTCGGGAC 402  
QY 419 AAATGCACTTGGCTCCACCCAGCTGCGCAAGAGTTTCTCATCTCGCCCTTCCTCC 478  
DB 403 AAGTCTTACTTACTGACCAACAGCCAGCAAGAGTTCTCTATCTCCCTCCGCGCTCA 462  
QY 479 CCACCTGTGCTGCGACCCCATCAAGCTGCGACGCTCCCACTATGACCTCCTC 538  
DB 463 CCCCCCGTGGGTGGAAGCAGAGTGAAGATGCCAGCCAGTGATCACTATGACCTCCTC 522  
QY 539 TATGCTGTGCCAACTAGGACAGGAGAGATGATGAGTCCATGAGGAGGAGTGC 598  
DB 523 TGTGCGCTCTCCAAAGCTGGGGCCAGGAGAGAAATACGAACCTGCACGCAACCGAGTCC 582  
QY 599 ACCCAAGTGTGCTGTCGACGTTGCCACAG 630  
DB 583 ACCCCAGTGTGCTGTCGACGTTGTGAGAG 614

RESULT 6  
US-08-665-040-1  
Sequence 1, Application US/08665040  
Patent No. 5869318  
GENERAL INFORMATION:  
APPLICANT: ESTIVILL PALLEJA, XAVIER  
APPLICANT: FUENTES, JUAN JOSE  
APPLICANT: PRITCHARD, MELANIE  
TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE  
TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION OF HUMAN CHROMOSOME 21,  
TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,  
TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY  
TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD  
TITLE OF INVENTION: FOR CHARACTERIZING IT.  
NUMBER OF SEQUENCES: 12

APPLICANT: Edwards, Carla M.  
APPLICANT: Streeter, David G.  
TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS  
FILE REFERENCE: PC-0013 US  
CURRENT APPLICATION NUMBER: US/09/614,474  
CURRENT FILING DATE: 2000-07-11  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PERL Program  
SEQ ID NO 8  
LENGTH: 615  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No. 6524819 206751\_Rn.1  
US-09-614-474-8

CORRESPONDENCE ADDRESS:  
ADDRESSEE: LADAS & PAREY  
STREET: 26 WEST 61ST STREET  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10023  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORD PERFECT 5.1 FOR DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,040  
FILING DATE: JUNE 7, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: ES P9501140  
FILING DATE: JUNE 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: JANET I. CORD  
REGISTRATION NUMBER: 33,778  
REFERENCE/DOCKET NUMBER: U010815-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 708-1800  
TELEFAX: (212) 246-8959  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2174 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: cDNA for mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: human  
ORGANISM: Homo sapiens  
DEVELOPMENTAL STAGE: foetal  
TISSUE TYPE: Brain  
IMMEDIATE SOURCE: gene library of cDNA  
LIBRARY: gene library of cDNA from foetal  
CLONE: BC-17.8-1 and BC-17.8-2  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9  
MAP POSITION: 21q22.1-q22.2  
FEATURE:  
NAME/KEY: cDNA for mRNA, BC-17.8  
LOCATION: 1..2174  
FEATURE:  
NAME/KEY: untranslated 5'  
LOCATION: 1..48  
FEATURE:  
NAME/KEY: coding sequence  
LOCATION: 49..560  
OTHER INFORMATION: Down Syndrome critical  
FEATURE:  
NAME/KEY: DSCR1  
LOCATION: 1..171 PEPTIDES  
IDENTIFICATION METHOD: translation of the  
OTHER INFORMATION:  
OTHER INFORMATION: - deduced protein  
OTHER INFORMATION: - proline-rich protein domains  
OTHER INFORMATION: - glutamic acid-rich protein domains  
OTHER INFORMATION: - leucine/phenylalanine-rich protein domains  
FEATURE:  
NAME/KEY: untranslated 3'  
LOCATION: 564..2174  
FEATURE:  
NAME/KEY: 2 poly (A)  
LOCATION: 1541..1546 AND 2132..2137  
US-08-665-040-1  
Query Match 19.0%; Score 177.4; DB 2; Length 2174;

RESULT 9  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.

us-10-030-613-3.rni

Tue Aug 26 09:11:05 2003

RESULT 10  
US-08-635-066-1  
Sequence 1, Application US/08635066  
Patent No. 5945580  
GENERAL INFORMATION:  
APPLICANT: Dunsuir, Pamela  
ADDRESSEE: Harpster, Mark H.  
TITLE OF INVENTION: Capsicum Hemiacellulase Polynucleotides  
TITLE OF INVENTION: and Polypeptides  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/635,066  
FILING DATE: 19-APR-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 012176-005500  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2254 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 609..2129  
US-08-635-066-1

Query Match 4.4%; Score 40.8; DB 1; Length 7218;  
Best Local Similarity 4.5%; Pred. No. 0.0096;  
Matches 18; Conservative 209; Mismatches 171; Indels 0; Gaps 0;  
393 TTCAGACTCCAGACAGATGGAGAACTGGCTTGGCTCCACCCAGCGCTGCCAAAC 452  
1031 TTCGAGCTGGCTGCGAGTGGAGGAGCTTGGCTGGCGAGCCATCAACGATGCCA 512  
453 AGTTCTCATCTCGCCCTTCTCCACCTGTTGGCTGGCGAGCCATCAACGATGCCA 512  
1091 YY 1150  
513 CGCAGPCCTCAACTATGCTCTGCTGGCCAACTAGGACCGAGGAGAGT 572  
1151 YY 1210  
573 ATGAGCTCATGAGGAGCTGAGTCCACCCAGTGTCTGCTGCGAGTGTGCGACAGT 632  
1211 YY 1270  
633 ACATAGAGGAG 692  
1271 YY 1330  
693 GTCCTGGCTGCCACCTCCGCTGCTCAACTGAGTGGCTGCTCTCTCGATAGCGG 752  
1331 YY 1390  
753 TCTCCTCTTATCATGCTTTTCCCTGTTGTTGTC 790  
1391 YY 1428  
US-08-232-463-14

RESULT 11  
US-09-328-352-3446  
Sequence 3446, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breston et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
Query Match 3.7%; Score 35; DB 2; Length 2254;  
Best Local Similarity 50.9%; Pred. No. 0.4;  
Matches 83; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
767 TCGTTTTCCCTCTGTTTGTCAAAAAAATTCCTTTAAATTCCTGGGTGTTGTT 826  
373 TGATTATGTCATATGTTTATCAATAAAAAATTTATGTTTTTTCATATTTTAT 432  
827 GTTTGAGATTCCTTCTGTTTATCAAGCTCTCGCAAAAAAGGCTAGAAAAAGTGATA 886  
433 GATTGCAATTTATTATGTTAAAAAATTTGGGAACAAATCTCTAATAGCAATA 492  
887 TGTCCTCTGATCATATCATACCATTAAGTATAACCATTTAT 929  
493 GCTACATGCCCTTAATAATACCAACCAATATCAAACTTT 535  
US-09-328-352-3446  
Sequence 3446, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breston et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352

```

; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3446
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3446

```

	Query Match	3.5%	Score 33;	DB 4;	Length 1011;
	Best Local Similarity	52.6%;	Pred. No. 1.2;		
	Matches	72;	Conservative	0;	Mismatches 65; Indels 0; Gaps 0;
QY	265	CCAGCTATTTAAGAGTTTCAGACGTGTCCTATATAAACTTCAGCAATCCTTAATCTGCAGC	324		
Db	252	CCGCTCGCTTTTAAAGTAAAAACCGTGCTTTATCAGCTTAAGCCAAAATTAATTCATCA	311		
QY	325	CCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAGAGGGAAAAATTAAGAGCTCTACT	384		
Db	312	ACTGCTATTAAAGAAATGATATGCACAAACTACCGCCCAAAAAGTTATTCAATACAT	371		
QY	365	TGCACAGGTCAGACTC	401		
Db	372	TGAACAAGATCACTC	388		

```

RESULT 12
US-09-141-000-4/c
; Sequence 4, Application US/09141000
; Patent No. 6054295
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING
; TITLE OF INVENTION: RECEPTOR PROTEINS
; FILE REFERENCE: 19999Y
; CURRENT APPLICATION NUMBER: US/09/141,000
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Human
US-09-141-000-4

```

```

Query Match      3.5%; Score 32.6; DB 3; Length 458;
Best Local Similarity 9.2%; Pred.No.0.97;
Matches 39; Conservative 125; Mismatches 251; Indels 0; Gaps 0;

Qy 228 AGGAGCTGTTTCGGACTTATCAGTCACGTGTGTGACGCTCCAGCTATTAAAGAGTTTCAGAC 287
    : : : | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 448 MDK.BMC..M.NR.D.BTMASA.Y...AK.KMCTYY.H.KD.CT.RH..T.D.BH..M. 389
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 288 GTGTCGGTATAAAGCTTCAGCACTCTAAATCTCGACGCCGAGCTAGGATAGAGCTTCATG 347
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 388 BT.BH..DKSHSNT.T.TM.AB...M.MKSMMEMB.....TNN.H..CT.MS.H.HK.RH 329
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 348 AAACCCAATTCAGAGGGAAAAAATTAAGCTCTACTTTGCACAGGTTTCAGACTCCAGAGA 407
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 328 HTRB..H..SYRBBC.....KWTs....SK.HT.S.AS.C..DMTWc..BB..YHT.HG 269
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 408 CAGATGGAGACAACACTGCATTTGGCTCCACCCCACCCTGCCAACACAGTTTCTCATCTCGC 467
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 .AA.TM.HSH.BA.H.KTR.M...TB..DSB.MNR.MMT..B...N.R.SN.TH.YYMRM 209
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 468 CCCCTTCCTCCCACTGTTGGCTGGCACCCCTCAACGATGCCACGCCAGCTCCTCAACT 527
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 208 .YCCBYH.YBC.M.KCBM.GMK.YGT.GSMYYMA...G.NAT.GSR..NC..AYM..TM 149
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Qy 528 ATGACCTCCCTCTATGCTGTGGCCAAACTTAGGACCCAGGAGGAAGTATGAGCTCCATGCAG 587
    | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
Db 148 GT.GSTBCRDRCST.HCSB.G.YM.KSN.KR.GMA..H.B.ASSGHR.MYB.CSCC... 89
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 588 GGACTGAGTGCCACCCCAAGTCGTCGTCGACGTGTCCGACAGTGACATAGAGG 641

```

[illegible]

```

>> LENGTH: 700
>> TYPE: DNA
>> ORGANISM: Homo sapiens
>> FEATURE:
>> NAME/KEY: misc_feature
>> LOCATION: (1)..(15652)
>> OTHER INFORMATION: k = A, T, G, or C
>> NAME/KEY: misc_feature
>> LOCATION: (1)..(15652)
>> OTHER INFORMATION: r = G or A; y = T/U or C; m = A or C;
>> OTHER INFORMATION: k = G or T/U; s = G or C; w = A or T/U;
>> OTHER INFORMATION: b = G, C, or T/U; d = A, G, or T/U;
>> OTHER INFORMATION: h = A, C, or T/U; v = A, G, or C
>> NAME/KEY: exon
>> LOCATION: (3419)...(3444)
>> NAME/KEY: intron
>> LOCATION: (3445)...(3908)
>> NAME/KEY: exon
>> LOCATION: (3909)...(3993)
>> NAME/KEY: intron
>> LOCATION: (3994)...(4694)
>> NAME/KEY: exon
>> LOCATION: (4695)...(4898)
>> NAME/KEY: intron
>> LOCATION: (4899)...(5652)
>> NAME/KEY: exon
>> LOCATION: (5653)...(5838)
>> NAME/KEY: intron
>> LOCATION: (5839)...(7184)
>> NAME/KEY: exon
>> LOCATION: (7185)...(7205)
>> NAME/KEY: intron
>> LOCATION: (7206)...(8310)
>> NAME/KEY: exon
>> LOCATION: (8311)...(8806)
>> NAME/KEY: intron
>> LOCATION: (8807)...(12271)
>> NAME/KEY: exon
>> LOCATION: (12272)...(12406)
>> NAME/KEY: intron
>> LOCATION: (12407)...(12820)
>> NAME/KEY: exon
>> LOCATION: (12821)...(12991)
>> NAME/KEY: intron
>> LOCATION: (12992)...(14089)
>> NAME/KEY: exon

```

us-10-030-613-3.rni

Tue Aug 26 09:11:05 2003

```

; LOCATION: (14090)...(14191)
; NAME/KEY: intron
; LOCATION: (14192)...(14477)
; NAME/KEY: exon
; LOCATION: (14478)...(14543)
; NAME/KEY: intron
; LOCATION: (14544)...(15002)
; NAME/KEY: exon
; LOCATION: (15003)...(15194)
; NAME/KEY: intron
; LOCATION: (15195)...(15652)
; US-09-422-936-60

```

```

Query Match 3.4%; Score 31.6; DB 4; Length 15652;
Best Local Similarity 53.2%; Pred. No. 17; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 59;

QY 648 AGGACCAAGAGCTTCCCAAGCCAAATATCATCCAAACTCGGGTCTGCGCTGCCAC 707
Db 9629 AAGTTCATAGGTGCCCCAGATCTCTATTATCCACGCTACCTCTTCCCTCTC 9570
QY 708 CCTCCGTGTCACAGTGCCTCTCTCTCGATAATAGCGCTCTCTTATCAT 767
Db 9569 CCCACTGCCATCTTTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9510
QY 768 GCTTTT 773
Db 9509 GCTATT 9504

```

```

RESULT 14
US-09-619-353-13
; Sequence 13, Application US/09619353
; Patent No. 6410249
; GENERAL INFORMATION:
; APPLICANT: Ngai, John
; APPLICANT: Speca, David J.
; APPLICANT: Lin, David M.
; APPLICANT: Isacoff, Ehud Y.
; APPLICANT: Dittman, Andrew H.
; APPLICANT: Fan, Jinhong
; TITLE OF INVENTION: Odorant Receptors
; FILE REFERENCE: B99-038-2
; CURRENT APPLICATION NUMBER: US/09/619,353
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/144,766
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2595
; TYPE: DNA
; ORGANISM: Brachydanio rerio (zebrafish)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)...(2592)
; US-09-619-353-13

```

```

Query Match 3.4%; Score 31.4; DB 4; Length 2595;
Best Local Similarity 54.9%; Pred. No. 6.9; Indels 0; Gaps 0;
Matches 62; Conservative 0; Mismatches 51;

QY 353 CAATTTCAGAGGAAATTAAGCTCTTCTTCCAGAGTTCAGACTCCAGAGACAGAT 412
Db 648 CATTACAGATGAGATTATGACGCTCTTCTTCCAGAGTTCAGACTCCAGAGACAGAT 707
QY 413 GGAGCAACTCTGCTTGGCTCCACCCAGCTCCGCAACAGTTTCTCATCTC 465
Db 708 GGAGGAATTTGATCGCTTTTAAAGCAATCTTCCAGACTCTCACTAGCAGATC 760

```

RESULT 15  
US-09-791-211-10

```

; Sequence 10, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791,211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 10
; LENGTH: 98844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 24962
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 64383
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65468
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65469
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65470
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65471
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 87130
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 89049
; OTHER INFORMATION: unknown
; OTHER INFORMATION:
; US-09-791-211-10

```

```

Query Match 3.4%; Score 31.4; DB 4; Length 98844;
Best Local Similarity 45.1%; Pred. No. 60; Indels 0; Gaps 0;
Matches 116; Conservative 0; Mismatches 141;

QY 121 AGAAGCCTTTCAAGCAATCACTGACTTCAATCACTCCCACTCGTTGTTGCGTCAA 180
Db 79522 AGACCCCATCTCTGAAAAAATGAGTGCATGTTACAGTGTCTTACAGTGTCTTA 79581
QY 181 TGTTCCAGTCACTGTTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 79582 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 79641
QY 241 GACTTATGATGACTGTGTGACGTTTCAGCTATTTAAGAGTTTCAGACGTTCCGTATA 300
Db 79642 GAGTGTGCTGCTACACCTGTATCCAGCAGCTCTGGAGGCGGAGCGGCGGATCATGA 79701
QY 301 CTTGAGCAATCTTAATCTCGACCGGAGTAGATAGAGCTTTCATGAACCCCAATTCAG 360
Db 79702 GGTTCAGGAGATCGAGACCATCTCTGGAGGCGGAGCGGCGGATCATCATGA 79761
QY 361 AGGGAATAATTAAGC 377
Db 79762 AAAAGAGAAAAATTAAGC 79778

```

Search completed: August 24, 2003, 12:13:07  
Job time : 100 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

```
Run on: August 22, 2003, 18:45:57 ; Search time 85 Seconds  
        (without alignments)  
        436.965 Million cell updates/sec
```

---

```
Title: US-10-030-613-I  
Perfect score: 1254  
Sequence: 1 MBSPCCCCCCTTTCCTTTCCGCGGGAGGATGCACGGTGCGGGGGG
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%

Database : A Geneset 1077702  
Listing first 45 summaries

```

1: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Length				
1	1254	100.0	234	22	AAB31788	-----	Amino acid sequenc
2	1254	100.0	255	23	AAW50760	-----	Human Down syndrom
3	913	72.8	192	23	AAE18916	-----	Human MCP associa
4	905	72.2	197	23	AAE18913	-----	Mouse MCP associa
5	729.5	58.2	241	23	AAE18917	-----	Human MCP associa
6	698.5	55.7	212	23	AAE18918	-----	Human MCP associa
7	629	50.2	197	23	AAE18914	-----	Human MCP associa
8	629	50.2	197	23	AAE18915	-----	Human MCP associa
9	623	49.7	198	23	AAE18912	-----	Mouse MCP associa

10	620.5	49.5	198	23	AAE18911
11	581	46.3	171	20	AAW73898
12	432.5	34.5	292	22	-
13	318	25.4	111	21	ABB71467
14	282	22.5	142	21	AAG01768
15	222	17.7	58	22	AAE58452
16	222	17.7	58	22	ABG49494
17	222	17.7	58	22	ABG51723
18	222	17.7	58	22	ABG29495
19	222	17.7	58	22	ABB34672
20	222	17.7	58	22	ABB36877
21	222	17.7	58	22	ABB20081
22	222	17.7	58	22	AAW55464
23	222	17.7	58	22	AAW57623
24	222	17.7	58	22	AAW67851
25	222	17.7	58	22	AAW15671
26	222	17.7	58	22	AAW17859
27	222	17.7	58	22	AAW28174
28	222	17.7	58	22	AAW30365
29	222	17.7	58	22	AAW40346
30	222	17.7	58	22	AAW05506
31	222	17.7	58	23	ABG37397
32	171	13.6	56	22	ABG39657
33	171	13.6	56	22	ABG48200
34	171	13.6	56	22	ABB28178
35	171	13.6	56	22	ABB33353
36	171	13.6	56	22	ABB19813
37	171	13.6	56	22	AAW54139
38	171	13.6	56	22	AAW65532
39	171	13.6	56	22	AAW14405
40	171	13.6	56	22	AAW26818
41	171	13.6	56	22	AAW02132
42	133	10.6	25	23	ABG36184
43	98	7.8	454	17	AAO08780
44	98	7.8	454	22	AAW97670
45	94	7.5	777	23	ABP41839

## ALIGNMENTS

[illegible]



QY 181 PGKYLHAGTSTPSVHVHVCSDIEEEEDPKTSPKPKIIQTRRPGLPSPSVN 234  
 DB 202 PGKYLHAGTSTPSVHVHVCSDIEEEEDPKTSPKPKIIQTRRPGLPSPSVN 255  
 RESULT 3  
 AAE18916  
 ID AAE18916 standard; Protein; 192 AA.  
 AC AAE18916;  
 XX XX  
 XX 17-MAY-2002 (first entry)  
 XX Human MCIP associated protein #3.  
 DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 XX heart failure; cardiomyopathy; heart disease; human.  
 KW Homo sapiens.  
 OS WO200204491-A2.  
 PN 17-JAN-2002.  
 PD 06-JUL-2001; 2001WO-US21662.  
 PF 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX Williams SR, Rothermel B;  
 PI WPI: 2002-179698/23.  
 DR N-PSDB; AAD30157.  
 XX Screening for modulators of muscle calcineurin interacting protein  
 (MCIP) binding, expression or phosphorylation, useful for treating  
 cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 calcineurin and a test compound -  
 XX  
 PS Example 1; Page 165-166; 174pp; English.  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated protein.  
 CC Note: This sequence has been described as human MCIP splice variant in  
 CC specification, however the sequence seems to be a MCIP associated  
 CC protein.  
 XX  
 SQ Sequence 192 AA;  
 Query Match 72.8%; Score 913; DB 23; Length 192;  
 Best Local Similarity 93.0%; Pred. No. 3.3e-89;  
 Matches 174; Conservative 3; Mismatches 10; Indels 0; Gaps 0;  
 QY 48 NSLFACNVHQSVEGSEKFKGLFTYDCCVTFQLEKSFRRVFNFSNPKSARARIE 107  
 DB 6 STLVACVVDVEFTNQVEKFKGLFTYDCCVTFQLEKSFRRVFNFSNPKSARARIE 65  
 QY 108 LHETQFRGKKLYFAQVOTPETDGDKHLAPQPAKQFLISPPSPVSWQPINATPV 167

Db 66 LHETQFRGKKLYFAQVOTPETDGDKHLAPQPAKQFLISPPSPVSWQPINATPV 125  
 QY 168 LNYDLLYAVAKLPGKEKYLHAGTSTPSVHVHVCSDIEEEEDPKTSPKPKIIQTRRPG 227  
 Db 126 LNYDLLYAVAKLPGKEKYLHAGTSTPSVHVHVCSDIEEEEDPKTSPKPKIIQTRRPG 185  
 QY 228 LPPSVSN 234  
 Db 186 LPPSVSN 192  
 RESULT 4  
 AAE18913  
 ID AAE18913 standard; Protein; 197 AA.  
 AC AAE18913;  
 XX XX  
 XX 17-MAY-2002 (first entry)  
 XX Mouse MCIP associated protein #4.  
 DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; mouse.  
 OS Mus musculus.  
 XX WO200204491-A2.  
 PN 17-JAN-2002.  
 PD 06-JUL-2001; 2001WO-US21662.  
 PF 07-JUL-2000; 2000US-216601P.  
 PR 13-FEB-2001; 2001US-0782953.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.  
 XX Williams SR, Rothermel B;  
 PI WPI: 2002-179698/23.  
 DR N-PSDB; AAD30154.  
 XX Screening for modulators of muscle calcineurin interacting protein  
 (MCIP) binding, expression or phosphorylation, useful for treating  
 cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 calcineurin and a test compound -  
 XX  
 PS Disclosure; Page 154; 174pp; English.  
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is mouse MCIP associated protein.  
 CC Note: This sequence has been described as mouse MCIP2 encoding DNA in  
 CC the specification, however the sequence seems to be a MCIP associated  
 CC protein.  
 XX  
 SQ Sequence 197 AA;  
 Query Match 72.2%; Score 905; DB 23; Length 197;  
 Best Local Similarity 91.4%; Pred. No. 2.4e-88;  
 Matches 171; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

us-10-030-613-1.1.rag

Tue Aug 26 09:11:01 2003

XX SQ Sequence 241 AA;  
 Query Match 58.2%; Score 729.5; DB 23; Length 241;  
 Best Local Similarity 66.1%; Pred. No. 2e-69;  
 Matches 144; Conservative 22; Mismatches 49; Indels 3; Gaps 2;  
 19 LCCIDRDWAVTRCFA--EEAQTATDNDLPNSLFACNVHQSVEGEEKFGLRTY 76  
 16 LCSTQDEEEEIFGENEDDDEMDLSDLPSTSLFACSVHEAVFEAREQERFEALFTY 75  
 77 DCVTFQLPKSFRRVRINFSNPKSAARARIELHETQFQKGLKLYFAQVQTPETDGLKH 136  
 76 DQVTFQLPKSFRRVRINFSNPKSAARARIELHETQFQKGLKLYFAQVQSGEVRDKSY 135  
 137 LAPPQAKOFLISPPSPVPVGNQPINATPVNLYLLYAVAKLGGKGYELHAGTESTPS 196  
 136 LLPQPVKOFLLISPPASPPVGNQSDAMPVINYDLLCAVSKLGGPKGYELHAGTESTPS 195  
 197 VVHVCDSDIEEEEDPKTSFKPKIOTRRPGLPPSVSN 234  
 196 VVHVCESETEEBETK-NPKQKIAQTRRRPPTAALN 232

RESULT 6:

AAE18918  
 ID AAE18918 standard; Protein: 212 AA.

XX AC AAE18918;  
 XX DT 17-MAY-2002 (first entry)

XX DE Human MCIP associated protein #5.  
 XX DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 XX KW heart failure; cardiomyopathy; heart disease; human.  
 XX OS Homo sapiens.

XX PN WO200204491-A2.  
 XX PD 17-JAN-2002.  
 XX PF 06-JUL-2001; 2001WO-US21662.  
 XX PR 07-JUL-2000; 2000US-216601P.  
 XX PR 13-FEB-2001; 2001US-0782953.  
 XX PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX PA (WILL/) WILLIAMS S R.  
 XX PA (ROTH/) ROTHERMEL B.

XX PI Williams SR, Rothermel B;  
 XX WPI: 2002-179698/23.  
 XX DR N-PSDB; AAD30159.  
 XX PT Screening for modulators of muscle calcineurin interacting protein  
 XX PT (MCIP) binding, expression or phosphorylation, useful for treating  
 XX PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 XX PT calcineurin and a test compound -  
 XX PS Disclosure; Page 171-172; 174pp; English.

XX CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 XX CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 XX CC complex with the catalytic subunit of calcineurin and increased levels  
 XX CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 XX CC transduction of certain target genes. The invention also relates to  
 XX CC methods for identifying modulators of MCIP binding, expression or  
 XX CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 XX CC antibodies and diseased tissues and subsequently for determining the  
 XX CC presence or absence of calcineurinopathy or as predictor of heart disease.  
 XX CC The present sequence is human MCIP associated protein.  
 XX CC Note: This sequence has been described as splice variant of MCIP1  
 XX CC initiated by exon 4 in the specification, however the sequence seems  
 XX CC to be a MCIP associated protein.

QY 48 NSLFACNVHQSVEGEEKFGLRTYDDCVTFQLPKSFRRVRINFSNPKSAARARIE 107  
 Db 11 STLIVACVVDVEFTNQEVKEKFEGLRTYDCVTFQLPKSFRRVRINFSNPKSAARARIE 70  
 QY 108 LHETQFQKGLKLYFAQVQTPETDGLKHLAPPQAKOFLISPPSPVPVGNQPINATPV 167  
 Db 71 LHETQFQKGLKLYFAQVQTPETDGLKHLAPPQAKOFLISPPSPVPVGNQPINATPV 130  
 QY 168 LNYDLLYAVAKLGGKGYELHAGTESTPSVVHVCDSDIEEEEDPKTSFKPKIOTRRPG 227  
 Db 131 LNTDLLYAVAKLGGKGYELHAGTESTPSVVHVCDSDIEEEEDPKTSFKPKIOTRRPG 190  
 QY 228 LPPSVSN 234  
 Db 191 LPPSVSN 197

RESULT 5

AAE18917  
 ID AAE18917 standard; Protein: 241 AA.

XX AC AAE18917;  
 XX DT 17-MAY-2002 (first entry)

XX DE Human MCIP associated protein #4.  
 XX DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 XX KW heart failure; cardiomyopathy; heart disease; human.  
 XX OS Homo sapiens.

XX PN WO200204491-A2.  
 XX PD 17-JAN-2002.  
 XX PF 06-JUL-2001; 2001WO-US21662.  
 XX PR 07-JUL-2000; 2000US-216601P.  
 XX PR 13-FEB-2001; 2001US-0782953.  
 XX PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX PA (WILL/) WILLIAMS S R.  
 XX PA (ROTH/) ROTHERMEL B.

XX PI Williams SR, Rothermel B;  
 XX WPI: 2002-179698/23.  
 XX DR N-PSDB; AAD30158.  
 XX PT Screening for modulators of muscle calcineurin interacting protein  
 XX PT (MCIP) binding, expression or phosphorylation, useful for treating  
 XX PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 XX PT calcineurin and a test compound -  
 XX PS Example 1; Page 168-169; 174pp; English.

XX CC The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 XX CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 XX CC complex with the catalytic subunit of calcineurin and increased levels  
 XX CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 XX CC transduction of certain target genes. The invention also relates to  
 XX CC methods for identifying modulators of MCIP binding, expression or  
 XX CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 XX CC antibodies and diseased tissues and subsequently for determining the  
 XX CC presence or absence of calcineurinopathy or as predictor of heart disease.  
 XX CC The present sequence is human MCIP associated protein.  
 XX CC Note: This sequence has been described as splice variant of MCIP1  
 XX CC initiated by exon 4 in the specification, however the sequence seems  
 XX CC to be a MCIP associated protein.

The invention relates to muscle calcineurin interacting proteins (MCIPs) and nucleic acid molecules encoding such proteins. MCIPs form a physical complex with the catalytic subunit of calcineurin. MCIPs form a physical complex with the catalytic subunit of calcineurin and increased levels of MCIPs correspond to a reduced ability of calcineurin to stimulate transcription of certain target genes. The invention also relates to

100

Example 1; Page 161-162; 174pp; English

Enclosure: page 151-152; 174pp; English.

Note: this sequence has been described as human MCIP3 in the present specification, however the sequence seems to be a MCIP associated protein.

	Score	Pred. NO.	Indels	Gaps
Very Match	49.7%	62.3	41	3
Best Local Similarity	62.4%	3.7e-58		
Matches	123	21	41	

SULT 10  
E18911

Human MCIP associated protein  
Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
Muscle calcineurin interacting protein; MCIP; human heart disease

FO200204491-A2.

2001WO-US21662.

07-JUL-2000; 2000US-216601P.

(TEXA ) UNIV TEXAS SYSTEM.

PA (ROTH/) ROTHENMEL D.

1

\_\_\_\_\_

New isolated nucleic acid detection reagent for detecting 1000 or more genes from *Drosophila* and for elucidating cell signalling and cell-cell

us-10-030-613-1.rag

Tue Aug 26 09:11:01 2003

PS Claim 13; SEQ ID 5849; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number

CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs

CC were prepared from total human RNAs or polyA+ RNAs derived from 30

CC different tissues. EST sequences usually correspond mainly to the 3'

CC untranslated region (UTR) of the mRNA because they are often obtained

CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for

CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in

CC those cases where longer cDNA sequences have been obtained, the full 5'

CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'

CC ends and can therefore be used to obtain full length cDNAs and genomic

CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and

CC chromosome mapping procedures. They are used to obtain upstream

CC regulatory sequences and to design expression and secretion vectors.

XX SQ Sequence 111 AA;

Query Match 25.4%; Score 318; DB 21; Length 111;

Best Local Similarity 57.7%; Pred. No. 6.9e-26;

Matches 64; Conservative 16; Mismatches 27; Indels 4; Gaps 1;

QY 42 DFNDLPNSLFACNVHQSVFEGESKEKFGFLRTYDDCVTFQLFKSFRRVRINFSNPKSA 101

DB 5 NFNFYSFSLIACVANSIDFSESETRAKFESLFRTYDKDITFQYFKSKVRINFSNPFSA 64

QY 102 ARARIELHETOFGRKKLKYFAQVQTPETDGLKHLAPPQPAKQFLISPPS 152

DB 65 ADARLQLHKTEPLGKEMKLYAQTLHGSS---HLAPPNPKQFLISPPA 111

RESULT 14

AAB58452

ID AAB58452 standard; Protein: 142 AA.

XX AC AAB58452;

XX 14-MAR-2001 (first entry)

XX Lung cancer associated polypeptide sequence SEQ ID 790.

DE Human; lung cancer associated protein; neuroprotective; cytostatic;

XX cardioactive; immunomodulatory; muscular active; vulnerary;

KW gastrointestinal; nephrotropic; antiinfective; gynecological;

KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;

KW proliferative disorder; wound healing; infectious disease.

XX OS Homo sapiens.

XX WO2000055180-A2.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05918.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Ruben SM;

XX WPI; 2000-587514/55.

XX N-PSDB; AAF18328.

XX Lung cancer associated gene sequences, referred to as lung cancer

PT antigens, useful for treatment, prevention, and diagnosis of disorders

PT such as lung cancer -

XX Claim 11; Page 1325-1326; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer

CC associated proteins represented in AAB58106 - AAB58548. Lung cancer

PT Interactions -

XX Disclosure; SEQ ID NO 41193; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 292 AA;

Query Match 34.5%; Score 432.5; DB 22; Length 292;

Best Local Similarity 41.1%; Pred. No. 1.5e-37;

Matches 92; Conservative 36; Mismatches 75; Indels 21; Gaps 6;

QY 4 PQQQGHVPEDGGLFLLCCIDRDWATRCFAEAFQAITDFNDLPNSLFACNVHQSVFEGE 63

DB 86 PNOHPSLPKEGD-----VDSD-----TEPEVDA-DSFDLPTSITVNIHSEVFANP 131

QY 64 ESKEKEGLFRTYDDCVTFQLFKSFRRVRINFSNPKSAARIELHETQFGRKK-LKLYF 122

DB 132 ELKAMEELFRFSESATFQWLRSPRLRVNYDNAIAANARIKHLQYEFNKKTIVTCYF 191

QY 123 AQVOTPETDGLKHLAPPQPAKQFLISPPSPVGPQINDATPVNLDLLYAVAKLGP 182

DB 192 AQVPTVPSN---KNLQPPAPVQFLISPPSPAGWEPREGEPLVNHDLAALSLTPG 248

QY 183 EKVHELHAGTSTPSVVVHVCDSIDIEEDPKTSPKPIITRRP 226

DB 249 ESHELHQSEDPALIVHTA---MLAETGFLQVQKAPIVQTKCP 289

RESULT 13

AAG01768

ID AAG01768 standard; Protein: 111 AA.

XX AC AAG01768;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 5849.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX gene therapy; chromosome mapping.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX N-PSDB; AAC01774.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for

XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for

PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

PT

XX



associated proteins and polynucleotide sequences, their agonists, and antagonists may have neuroprotective; cytotatic; cardioactive; and immunomodulatory; muscular active general; vulnerable; gastrointestinal general; nephrotropic; antineoplastic; gynecological; or antibacterial activity. The invention also includes antibodies specific for the protein or polynucleotide sequences. The lung cancer associated protein or polynucleotide sequences may be used for detection of lung cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The proteins may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders. The proteins may also be used in the treatment of wounds and infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and peptide AAH5849 are used in the course of the invention for the identification and characterisation of the polynucleotide and protein sequences.

SQ Sequence 142 AA:

```

Query Match      22.5%; Score 282; DB 21; Length 142;
Best Local Similarity 46.7%; Pred. No. 7e-22;
Matches 63; Conservative 15; Mismatches 39; Indels 18; Gaps 3;

12 EDGGLLCCIDRDNAVTRCFEAEEFQAITFDNLPSLFCACNVHQSVFGESEKKEFG 71
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
26 EGG-----DWSXIDCEEE-----VDQLDPSATIACHLDPRFVFDGLCKRAKFFS 71
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
72 LFRTYDDCVTFOLFKSFRVRINFSNPKSAAARIELHETQFRGKKLKYFAQVQTPETD 131
   ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72 LFRTYDDKIDITQYKFSKFRVRINFSNPFSAADARQLLHKTEFLCKEMKLYFAQTLHGSS 131
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 GDKLHLAPQPAKQF 146
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 ----HLAPQIQTSSF 142

```

RESULT 15

ABG49494.			
ID	ABG49494	standard; Peptide; 58 AA.	
XX	AC	ABG49494;	
XX	DT	25-FEB-2003	(first entry)
XX	DE	Human liver peptide, SEQ ID No 28142.	
XX	DE	Human liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia; coronary heart disease.	
KW	OS	Homo sapiens.	
PN	WO	2000157273-A2.	
PD	09-AUG-2001.		
PF	30-JAN-2001; 2001WO-US00664.		
PR	04-FEB-2000; 2000US-0180312.		
PR	26-MAY-2000; 2000US-0207456.		
PR	30-JUN-2000; 2000US-0608408.		
PR	03-AUG-2000; 2000US-0632366.		
PR	21-SEP-2000; 2000US-0234687.		
PR	27-SEP-2000; 2000US-0236359.		
PR	04-OCT-2000; 2000GB-0024263.		
XX	XX		

PA  
XX (MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

DR  
yy  
WPI; 2001-488898/53.

Human genome-derived single exon nucleic acid probes useful for analysing gene expression in human adult liver -

XX

Claim 27; SEQ ID No 28142; 658pp; English.

The invention relates to a single exon nucleic acid probe (SENP) (I) for measuring human gene expression in a sample derived from human adult liver, comprising one of 13109 defined nucleotide sequences given in the specification (or complements/ fragments). The probe hybridises at high stringency to a nucleic acid molecule expressed in the human adult liver. (I) may be used for predicting, measuring and displaying gene expression in samples derived from human adult liver. The genes identified may be involved in genetic liver diseases such as cirrhosis, hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is associated with coronary heart disease. ABG47348-ABG59930 represent human liver single exon encoded peptides of the invention.

Note: The sequence information for this patent does not appear in the printed specification but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

Sequence 58 AA:

### Query Match

Best Local Matches	Similarity	Score	DB	Length
42; Conservative	73.7%	222;	22;	58;
		Pred. No. 5.1e-16;		
		8; Mismatches	7; Indels	0; Gaps
QY	68	KFEGLFRYYDCVTFQLEKSRVRRVINSNPKSAAARIELHETQFGKGLKLYFAQ		
DB	2	KFEGLFRYYDCVTFQLEKSRVRRVINSNPKSAAARIELHETQFGKGLKLYFAQ		124

Search completed: August 22, 2003, 18:56:09  
Job time : 87 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 18:53:37 ; Search time 40 seconds  
(without alignments)  
562.587 Million cell updates/sec

Title: US-10-030-613-1

Perfect score: 1254

Sequence: 1 MRSPGQQGHVPEDGGLFLC.....SPKXIIQTRRGLPPSYSN 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	351.5	28.0	239	T34305	hypothetical prote
2	146	11.6	315	T49641	nebula related pro
3	107.5	8.6	163	S62444	conserved hypothet
4	98	7.8	454	A56954	yes-associated pro
5	96	7.7	277	T37629	hypothetical prote
6	94	7.5	813	A72203	cellobiose-phospho
7	93.5	7.5	751	AC2098	hypothetical prote
8	92	7.3	270	T19033	hypothetical prote
9	89	7.1	3329	T42205	breast cancer susc
10	88	7.0	3329	T30904	breast cancer tumo
11	87.5	7.0	1557	T13160	protein CNK - frui
12	87	6.9	671	C96534	probable Poly-A Bi
13	86.5	6.9	1331	A48954	mannan endo-1,4-be
14	86	6.9	668	B96740	hypothetical prote
15	85	6.8	1232	S40766	hypothetical prote
16	84.5	6.7	212	T48580	hypothetical prote
17	84.5	6.7	1175	T22491	hypothetical prote
18	84.5	6.7	3938	T42761	atrophin-1 related
19	84	6.7	1006	T42731	formin isoform IV
20	84	6.7	1206	S24407	formin - mouse
21	84	6.7	1468	S11515	protein Y41E3.7 [i
22	83.5	6.7	338	A88908	hypothetical prote
23	83.5	6.7	435	T26818	hypothetical prote
24	83	6.6	160	C72706	hypothetical prote
25	83	6.6	643	T23453	hypothetical prote
26	83	6.6	3328	T30835	breast cancer tumo
27	82.5	6.6	457	S03961	Ig mu chain C regi
28	82.5	6.6	663	T46611	Mx protein - pig
29	82.5	6.6	1201	T29329	hypothetical prote

## ALIGNMENTS

### RESULT 1

T34305

hypothetical protein F54E7.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T34305

R:Bentley, D.

submitted to the EMBL Data Library, June 1994

A:Description: The sequence of C. elegans cosmid F54E7.

A:Reference number: 221502

A:Accession: T34305

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-239 <BEN>

A:Cross-references: EMBL:U00067; PIDN:AAC77519.1; GSPDB:GN00021; CESP:F54E7.7

A:Experimental source: strain Bristol N2; clone F54E7

C:Genetics:

A:Gene: CESP:F54E7.7

A:Map position: 3

A:Introns: 40/3; 78/2; 165/2; 205/1

Query Match	28.0%	Score 351.5;	DB 2;	Length 239;
Best Local Similarity	39.3%;	Pred. No. 6.5e-24;		
Matches	72;	Conservative	38;	Mismatches 66;
				Indels 7;
				Gaps 4;
Qy	40	ITDNDLPNSLFAACNVHQSVEGEGESKEFGLRTYDDCVTFQLFKSRFRVFNPNK	99	
Db	22	VSSKDDLPAIIVTQVPEDVDNKNQKAFSSSLFTQTEKDIHFDLRSFRVRVIFSSPE	81	
Qy	100	SAARARTELHETOPRGKKLKYFAOVTPTDGDGKLHLAPPQAKQFLISPPSPVGVQW	159	
Db	82	NATAKLIIVOGFSFGHELRKAFFA--ORIYMSANSQMLSPPLEKQFLISPPSPVGVWE	139	
Qy	160	PINDATPVL-NYDLLAYAVAKLGPGEKVELHAGTESTPSVVVHVCDSDIEEE---EDPKTS	215	
Db	140	QTKDMPVVCNEDLMARLASFAIDKEVEVNGDELTPAIIIVHPCETPIDVPSAIEMPRT-	198	
Qy	216	PKP 218		
Db	199	PRP 201		

### RESULT 2

T49641

nebula related protein [imported] - Neurospora crassa

N:Alternate names: protein B5022.230

C:Species: Neurospora crassa

C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C:Accession: T49641

R:Schulte, U.; Align, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu

submitted to the Protein Sequence Database, May 2000

A:Reference number: 225022

A:Accession: T49641  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-315 <SCH>  
 A:Cross-references: EMBL:AL355932; GSPDB:GN00116; NCSP:B5022.230  
 A:Experimental source: BAC clone B5022; strain OR74A  
 C:Genetics:  
 A:Gene: NCSP:B5022.230  
 A:Map position: 6

Query Match 11.6%; Score 146; DB 2; Length 315;  
 Best Local Similarity 31.0%; Pred. No. 2e-05;  
 Matches 49; Conservative 21; Mismatches 60; Indels 28; Gaps 7;

QY 42 DFNDLP-----NSLFACNVHQ-SVFEGESKEFGLFRYDDCV--TFOLF---- 85  
 DB DUSNUPPQSPPPPNTLFTNSLDFVSADN-----LQIRDLISOTAFIFAWSP 99  
 QY 86 -KSFRVRINFNSPKSAARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQPAK 144  
 DB 100 LKSFRIITVFDEQAALAVRSVMDGEAILGRCRVYFGQPTPIDVSAADKHLALPDAGK 159  
 QY 145 QFLISPPSPVGVW-OPINDA--TPVLYNDLLYAVAKL 179  
 DB 160 LFFISPPSPPHDWSQRMEDAPNTVMHAEADLAELAKL 197

## RESULT 3

S62444 conserved hypothetical protein SPAC13G6.15c - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: hypothetical protein SPAC24B11.04c  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 10-Dec-1999  
 A:Accession: S62444; S62549; T37650; T38331  
 R:Odell, C.; Bowman, S.  
 Submitted to the EMBL Data Library, October 1995  
 A:Reference number: S62430  
 A:Accession: S62444  
 A:Molecule type: DNA  
 A:Residues: 1-163 <ODE>  
 A:Cross-references: EMBL:Z54308; NID:g1008985; PIDN:CAA91108.1; PID:g1009000  
 R:Odell, C.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.  
 Submitted to the EMBL Data Library, October 1995  
 A:Reference number: Z21734  
 A:Accession: T37650  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-163 <OD3>  
 A:Cross-references: EMBL:Z54308; PIDN:CAA91108.1; GSPDB:GN000066; SPDB:SPAC13G6.15c  
 R:Odell, C.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 Submitted to the EMBL Data Library, November 1995  
 A:Reference number: Z21786  
 A:Accession: T38331  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-163 <OD4>  
 A:Cross-references: EMBL:Z67757; PIDN:CAA91769.1; GSPDB:GN000066; SPDB:SPAC24B11.04c  
 A:Experimental source: strain 972h-; cosmid c24B11  
 C:Genetics:  
 A:Gene: SPAC24B11.04c; SPAC13G6.15c; SPAC24B11.04c  
 A:Map position: 1L  
 C:Superfamily: fission yeast hypothetical protein SPAC13G6.15c

Query Match 8.6%; Score 107.5; DB 2; Length 163;  
 Best Local Similarity 34.9%; Pred. No. 0.024;

Matches 29; Conservative 13; Mismatches 28; Indels 13; Gaps 3;  
 QY 137 LAPPQAKOFLISPPSPVGVQWOPINDATP---VLNVDLLYAVAKLGPGEKVELHAGTES 193  
 DB 85 LQVPKEKNWLISSPPSPVGVQWOPINDATP---VLNVDLLYAVAKLGPGEKVELHAGTES 141  
 QY 194 TPSVVHVCDSDIEEEDPKTSP 216  
 DB 142 -POIVIS-----EHNNTKETSP 157

## RESULT 4

A56954 yes-associated protein, 65K - human  
 N:Alternate names: yap65 protein  
 C:Species: Homo sapiens (man)  
 C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 23-Sep-2002  
 A:Accession: A56954; S46973  
 R:Sudol, M.; Bork, P.; Einbond, A.; Kastury, K.; Druck, T.; Negrini, M.; Huebner, K.; J. Biol. Chem. 270, 14733-14741, 1995  
 A:Title: Characterization of the mammalian YAP (Yes-associated protein) gene and its  
 A:Reference number: A56954; MUID:95301570; PMID:7782338  
 A:Accession: A56954  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-454 <SUD>  
 A:Cross-references: EMBL:X80507; NID:g517176; PIDN:CAA56672.1; PID:g517177  
 C:Genetics:  
 A:Gene: YAP  
 C:Superfamily: yes-associated protein; WW repeat homology  
 F:171-208/Domain: WW repeat homology <WW1>

Query Match 7.8%; Score 98; DB 2; Length 454;  
 Best Local Similarity 31.1%; Pred. No. 0.62;  
 Matches 33; Conservative 11; Mismatches 30; Indels 32; Gaps 5;  
 QY 139 PPQPAKQFLISPPSPVGVQWOPINDATPVLNVDLLYAVAKLGPGEKVELHAGTESPSV- 197  
 DB 8 PPQPAQGGQGGPPSPVGVQWOPINDATPVLNVDLLYAVAKLGPGEKVELHAGTESPSV- 48  
 QY 198 -----VVHV-CDSDIEEEE-----DPKSPKPKIITRRPGLPPS 231  
 DB 49 PACHQIVHVRGDSSETDLFNAVMNPKTANVQTPVPMRLKLPDS 94

## RESULT 5

T37629 hypothetical protein SPAC13F5.04c - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T37629  
 R:Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 Submitted to the EMBL Data Library, September 1997  
 A:Reference number: Z21733  
 A:Accession: T37629  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-277 <BRO>  
 A:Cross-references: EMBL:Z99091; PIDN:CAB11767.1; GSPDB:GN000066; SPDB:SPAC13F5.04c  
 A:Experimental source: strain 972h-; cosmid c13F5  
 C:Genetics:  
 A:Gene: SPDB:SPAC13F5.04c  
 A:Map position: 1  
 A:Introns: 20/3

Query Match 7.7%; Score 96; DB 2; Length 277;  
 Best Local Similarity 22.4%; Pred. No. 0.5;  
 Matches 46; Conservative 22; Mismatches 77; Indels 60; Gaps 7;

QY 79 CVTFQLFKSFRVRINF-----SNPKSAAR-----ARIELHETQFRGK 116  
 DB 2 CFFPRLYSSFAILRISFLTSIKFKCKLRALQVANPQKAKTPSNHATEELOQSSNST 61

QY 117 KLKLYFAOVQTPETDGDKLHLAPP--QPAKQFLISPPSPVGVQWP--INDA 164  
 Db 62 TLPTQEAARVETNASHETSFALPTTSPAASLSISPTKSAVSEPNVADVKLSSTPA 121  
 QY 165 TPVLNLYLLYAVAKLGPEKYE-----LHAGTSTPSV-----VVHVCDSIDEE 208  
 Db 122 APQLN-----SPSHSYETPTSTSTITENLPTIDTRSTRSSHIQSLSPESKQ 171  
 QY 209 EEDPKTSPKPKIITRRPGLPPSVS 233  
 Db 172 TSDGHRPPSPSTITTTSTSIDPSVA 196

RESULT 6  
 A72203  
 cellobiose-phosphorylase - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: A72203  
 R:Neilson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
 A:Reference number: A72200; PMID:99287316; PMID:10360571  
 A:Accession: A72203  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-813 <ARN>  
 A:Cross-references: GB:AE001822; GB:AE000512; NID:g4982429; PIDN:AD36910.1; PID:g498243  
 A:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: TM1848

Query Match  
 Best Local Similarity 7.5%; Score 94; DB 2; Length 813;  
 Matches 45; Conservative 27; Mismatches 78; Indels 34; Gaps 9;

QY 51 FACNVHQSVEGEESKEFEGFLFRPYDCVTFQL--FKSPRRVRINFSNPKSAARARIEL 9;  
 Db 126 FTGEVHVLVLENAEKAPKIKLFSFIEFLNALLDDMTNQR---NYSTGEVIEGSIY 108  
 QY 109 HETQTRGKKLYFAOVQTP-----ETDGDK-LHLAPPQPAKQFLI--SPSPSPVGVQWPI 182  
 Db 183 HKTEYRRNRHYAFVSNQPIDGFDTRDSFGLYSGFEAPQAVVEGKPRNSVASGAP 161  
 QY 162 NDATPVNLNLYLLYAVAKLGPEKYEELHAGTSTPSVVHVCDSIDEEEDPKTSPKPKII 242  
 Db 243 ASH-----YLEIELAPSEKKEL-----IFIL-CYVENPEEERWE-KPGVI 221  
 QY 222 QTRR 225  
 Db 281 NKRR 284

RESULT 7  
 AC2098  
 hypothetical protein alr2338 [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp. strain PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C:Accession: AC2098  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
 A:Reference number: AB1807; PMID:21595285; PMID:11759840  
 C:Accession: AC2098  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-751 <KUR>  
 A:Cross-references: GB:BA0000019; PIDN:BA074037.1; PID:g17131430; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120

C:Genetics:  
 A:Gene: alr2338

Query Match  
 Best Local Similarity 7.5%; Score 93.5; DB 2; Length 751;  
 Matches 48; Conservative 30; Mismatches 95; Indels 45; Gaps 7;

QY 44 NDLPNSLFACNVHQSVEGEESKEFEGFLFRPYDCVTFQLFKSPRRVRINFSNPKSAAR 103  
 Db 109 NOTESQIAALKSAQSIGPAIENNEQVEKMLGSDLDNLTI-AFNSLQQLKKNYQNSLSQL 167  
 QY 104 ARIELHETO-----FRGKKLYFAOVQTPETDGDKLHLAPPQPAKQFLISPPS 152  
 Db 168 SRMYNQOQGETIVEELIDRLRGELT-----AIQETSTAKAQLSPPTVLOPPELOPPS 221  
 QY 153 SPV-----GWOPINDATPVNLNLYLLYAVAKLGPEKYEELHAGTSTPSV 197  
 Db 222 SPVVNLSPPTVLOFPDQSQSPNLAQSTPLEETSTTKPSVSITPPEK-----STPVT 273  
 QY 198 VHVCDSDIEEED--PKTSP--KPKIQTTRRGLPPS 231  
 Db 274 IVPPPQETRRPETKSVIPKVPSPDSETKLOSQOEKAAEPS 311

RESULT 8  
 T19033  
 hypothetical protein C07A4.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T19033  
 R:White, S.  
 submitted to the EMBL Data Library, August 1996  
 A:Reference number: Z19063  
 A:Accession: T19033  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-270 <WIL>  
 A:Cross-references: EMBL:Z78536; PIDN:CAB01717.1; GSPDB:GN00028; CESP:C07A4.1  
 A:Experimental source: clone C07A4  
 C:Genetics:  
 A:Gene: CESP:C07A4.1  
 A:Map position: X  
 A:Introns: 54/2; 79/3; 119/3; 142/1

Query Match  
 Best Local Similarity 7.3%; Score 92; DB 2; Length 270;  
 Matches 42; Conservative 31; Mismatches 72; Indels 26; Gaps 8;

QY 25 DWAVTRCFABEAFOAITDNDLPNSLFACNVHQSVEGEESKEFEG-----LFRTYDDCV 80  
 Db 101 NWA-ARKNSEENRDKLT-FEQVFNSTKADNT--SVYVGNISQOTTDDADLRDLFSYGDIA 156  
 QY 81 TFOLEKSFRRVRINFSNPKSAARARIELHETQFRCKKLYFAOVQTPETDGD-----133  
 Db 157 EVRIFQRYAFVRYEKKECATKALMENKEMAGNQCVRCSWGRTOAVVNNQLSNALKI 216  
 QY 134 -KLHLAP-----PQPAKQFLISPPSPVGVQWOPINDATPVNLNYD---LLYA 175  
 Db 217 TVFNLPQNALNPLPIDLSLMLMPTLPTIPLION--PFLNVEPATLLYS 265

RESULT 9  
 T42205  
 breast cancer susceptibility protein BRCA2 - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
 C:Accession: T42205  
 R:McAllister, K.A.; Haugen-Strano, A.; Hagevik, S.; Collins, N.K.; Brownlee, H.; Futr  
 submitted to the EMBL Data Library, February 1997  
 A:Description: Characterization of the mouse and rat homologs of the BRCA2 breast can  
 A:Reference number: Z22073  
 A:Accession: T42205  
 A>Status: preliminary; translated from GB/EMBL/DBJ





entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; AF282255; AAF91461.1;
CC EMBL; AF237789; AAF63485.1;
CC EMBL; AF237790; AAF63486.1;
CC EMBL; AF260717; AAF70343.1;
CC EMBL; AF263239; AAF72701.1;
CC EMBL; AF263240; AAF72702.1;
CC EMBL; AK010696; BAB27128.1;
CC EMBL; BC013551; AAH13551.1;
CC MGD; MGI:1890564; Dscr1.
CC GO; GO:0005737; C:cytoplasm; IDA.
CC GO; GO:0005634; C:nucleus; IDA.
CC InterPro; IPR006931; Calcipressin.
CC Pfam; PF04847; Calcipressin; 1.
KW Alternative splicing.
FT VARSPLIC 1 28 MEEVDLQPLSATIACHLDPFRVFDGLC -> MHERDFSYN
FT FSSLIACVANDVFESEET (in isoform B).
FT /FTID=VSP_001317.
FT Missing (in isoform C).
FT /FTID=VSP_001318.
FT T -> P (IN REF. 2).
FT CONFLICT 152 152
FT SEQUENCE 198 AA; 22706 MW; BA35340AFD6F0582 CRC64;

Query Match 49.6%; Score 622.5; DB 1; Length 198;
Best Local Similarity 62.4%; Pred. No. 1.7e-46;
Matches 121; Conservative 23; Mismatches 41; Indels 9; Gaps 2;

QY 42 DFNPLNSLFCNQHVSFEGESKEPEGLFRTYDVCVTFQFLKSPRRVIRNFSNPKSA 101
DB 5 DLQPLSATIACHLDPFRVFDGLCRKPFSLFRTYDVTTFQYKPKRVIRNFSNPLSA 64
QY 102 ARARIELHTQPRGKKLLKLYFAQVQTPETDGDKLHLAPPKAFOLISPPSSPPVGWQPI 161
DB 65 ADARLRHLTEFLGKEMKLYFAQTLHGSS----HLAPPNDKQFLISPPASPPVGWQV 120
QY 162 NDATPVNLVLDLYAVAKLPGKYLHAGTSTPTSVVHVHVCDSIDEEEDPK-----TSP 216
DB 121 EDATPVNLVLDLYAVAKLPGKYLHAGTSTPTSVVHVHVCDSIDEEEDPK-----TSP 216
QY 217 KPRIIOTRRRGLPP 230
DB 181 KPRIIOTRRRPEXT 194

RESULT 8
ID SRA_DROME STANDARD; PRT; 292 AA.
AC Q9XZL8; Q9V391;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sarah protein (Nebula protein).
GN SRA OR NLA OR CG6072.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA McCormick A.V., Goldberg M.L.;
RT "Gene required for elongation of meiosis I spindle in Drosophila females."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

```

```

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari P., Bottier P.,
RA Borkova D., Botchan M.R., Bouck J., Brodeur E., Center A., Chandra I.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Caceres L., Davies P.,
RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabriellian A.E., Garg N.S., Gu Z., Guan P., Harris M.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Hernandez J.R., Houck J.,
RA Harris N.L., Harvey D., Heiman T.J., Wei M.-H., Ibegwam C.,
RA Hostin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Poillard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
Science 287:2185-2195(2000).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN-Berkeley;
RC MEDLINE=22426066; PubMed=12537569;
RX Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarini H., Kronmiller B., Pacleeb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource."
RL Genome Biol. 3:RESEARCH080.1-RESEARCH080.8(2002).
CC -!- FUNCTION: REQUIRED FOR ELONGATION OF MEIOSIS I SPINDLE.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF147700; AAD33987.1;
CC EMBL; AE003712; AAF55285.1;
CC EMBL; AY061194; AAL28742.1;
CC FlyBase; FBgn0020250; sra.
CC InterPro; IPR006931; Calcipressin.
CC Pfam; PF04847; Calcipressin; 1.
CC SEQUENCE 292 AA; 31423 MW; 64F1BBF5F6EA6CF9 CRC64;

Query Match 34.5%; Score 432.5; DB 1; Length 292;
Best Local Similarity 41.1%; Pred. No. 5.7e-30;
Matches 92; Conservative 36; Mismatches 75; Indels 21; Gaps 6;

QY 4 PGQGHVPEPGLGLCCIDRDWAVTRCFEAFAQITDFNDPLNSLFCNQHVSFEGE 63
DB 86 PNQHPSLPKREGD-----VDSD-----TEPEVDA-DSFDPLPISIIIVNIHSEFANP 131
QY 64 ESKEKEPGLFRTYDVCVTFQFLKSPRRVIRNFSNPKSAARTELETQFGRKK-LKLYF 122

```



Db 132 ELKAMEELRFTSESATFQWLSFRRLRVNYDNAIAAANARIKLHOYEENKKTITCVF 191  
 QY 123 AQVOTPETDGLHLPQPAKQFLISPPSPVGPQINDATPVNLDLLYAVAKLPG 182  
 Db 192 AQVTPVSN--KNLOPPAPVKQFLISPPSPAGWEPREEPELVNHDLLAALSLTPG 248  
 QY 183 EYKELHAGTESPVVVHVCDSDIEEDPKSPKPKIQTTRP 226  
 Db 249 ESHLHPOSQDQAIIVHTA---MLAETGGLQVRAPIVOTKCP 289

## RESULT 9

CCPL\_CAREL STANDARD; PRT; 207 AA.  
 AC P53806: Q9U6V5;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Calciopressin-like protein (Down Syndrome candidate region 1-like protein).  
 DE RCN-1 OR DSCRIL OR F54E7.7.  
 GN Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Petoderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Bentley D.;  
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=2021370; PubMed=10756093;  
 RA Strippoli P., Lenzi L., Petri M., Carinci P., Zannotti M.;  
 ET "A new gene family including DSCR1 (Down syndrome candidate region 1) and Zaki-4: characterization from yeast to human and identification of DSCR1-like 2, a novel human member (DSCRIL2).";  
 RT Genomics 64:253-263(2000).  
 RL Genomics 64:253-263(2000).  
 CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses by binding to the catalytic domain of calcineurin A (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U00067; AAK20076.1; -  
 DR EMBL; AF176115; AAF01683.1; -  
 DR PIR; T34305; T34305.  
 DR WormPep; F54E7.7; CE01318  
 DR InterPro; IPR006931; Calciopressin.  
 DR Pfam; PF04847; Calciopressin; 1.  
 FT CONFLICT 205 207  
 FT CEQ -> SLQSSSFVLCCTPSPFFSPFPCFNFNPQSLFCR  
 G (IN REF. 1)  
 SQ SEQUENCE 207 AA; 23030 MW; 0154E308AB05B79 CRC64;

Query Match  
 Best Local Similarity 28.0%; Score 351.5; DB 1; Length 207;  
 Matches 72; Conservative 38; Mismatches 66; Indels 7; Gaps 4;  
 QY 40 ITDFNDLNSLPACNVHVSFEGESKEKFEGLFTYDCVTFOLFSPFRVRINFSNPK 99  
 Db 22 VSSKDDLPNAIVTQVPDVEDFNQDKRANFSSLTQIEKDIHDFDLNFSRRVRVIFSSPE 81  
 QY 100 SAARIELHETQFGKLLKLYFAQVOTPETDGLHLPQPAKQFLISPPSPVGPQINDATPVNLDLLYAVAKLPG 182  
 Db 82 NATAKLIVQFSFKGHEHKAFFA--ORIYMSANSOMLSPPLEKQLISPPCSPVPGWE 139

QY 160 PINDATPVL-NYDLLYAVAKLPGERYELHAGTESPVVVHVCDSDIEE---EDPKTS 215  
 Db 140 QTKMDPPVVCNFDLMARLAFIDEKYEYHNGDELTPALIVHPCTPIDVPSAIEPRT- 198  
 QY 216 PKP 218  
 Db 199 PRP 201

## RESULT 10

YA9F\_SCHPO STANDARD; PRT; 163 AA.  
 AC Q09791;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C13G6.15c in chromosome I.  
 GN SPAC13G6.15c OR SPAC24B11.04C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Church C.M., Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., McManus P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E., Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B., Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RL "The genome sequence of Schizosaccharomyces pombe".  
 CC Nature 415:871-880(2002).  
 CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses by binding to the catalytic domain of calcineurin (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; Z54308; CAA91108.1; -  
 DR EMBL; Z67757; CAA91769.1; -  
 DR PIR; S62444; S62444  
 DR GeneDB.SPombe; SPAC13G6.15c; -  
 DR InterPro; IPR000931; Calciopressin.  
 DR Pfam; PF04847; Calciopressin; 1.  
 DR Hypothetical protein.

```

171 204      WW.
DOMAIN 171 204      WW.
SEQUENCE 454 AA; 48755 MW; 87CB840D3393EFC0 CRC64;

Query Match      7.8%; Score 98; DB 1; Length 454;
Best Local Similarity 31.1%; Pred. NO. 0.49;
Matches 33; Conservative 11; Mismatches 30; Indels 32; Gaps 5;

139 PPQAPKFLISPPSSPPVGQNPINDATPVLTDLLYAVAKLGPGEYELHAGTSTPSV- 197
      ||||| ||||| ||||| ||||| |||||
      8 PPQAPAPQGGPPSPQPGGPPS-----GPGQ--PAPATQAQPAQ 48

198 -----VVHV-CDSDIEEE-----DPKTSPPKLIOTRRPGLPPS 231
      ||||| ||||| ||||| ||||| |||||
49 PAGHOIVHVRGDSFETDLEALFNAMNPKTANVPQTPMRLKLPDS 94
      ||||| ||||| ||||| ||||| |||||
db

```

## RESULT 12

```

ID      NEOL_RAT      STANDARD      PRPT; 1377 AA.
AC      P97603;
AD      16-OCT-2001 (Rel. 40, Created)
AD      16-OCT-2001 (Rel. 40, Last sequence update)
DT      DT
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DT      DT
DT      Neogenin precursor (Fragment).
DE      NEOL OR NGN.
GN      Rattus norvegicus (Rat)..
OS      Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Eukaryota; Metazoa; Chordata; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC      Mammalia; Eutheria;
OC      NCBI_TaxId=10116;
OX      [1]
RN      SEQUENCE FROM N.A.
RP      TISSUE=Brain;
RC      MEDLINE=97015074; PubMed=8861902;
RX      Keino-Masu K., Masu M., Hinck L., Leonardo E.D., Chan S.S.-Y.,
RA      Culotti J.G., Tessier-Lavigne M.;
RA      Deleted in Colorectal Cancer (DCC) encodes a netrin receptor.";
RT      Cell 87:175-185(1996).
RL      -1- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE
CC      TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
CC      DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
CC      MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC      SUBFAMILY.
CC      -1- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.
CC      -1- SIMILARITY: Contains 6 fibronectin type III domains.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/).
```

CC	or send an email to <a href="mailto:license@cs.cmu.edu">license@cs.cmu.edu</a>	
CC	EMBL; U68726; AAB41100.1; "	
DR	HSSP; P56276; ITLK.	
DR	InterPro: IPR003961; FN.III.	
DR	InterPro: IPR003962; FN.III.Subd.	
DR	InterPro: IPR007110; I9-like.	
DR	InterPro: IPR003598; I9-C2.	
DR	InterPro: IPR003006; I9_MHC.	
DR	Pfam: PF00041; fn3; 6.	
DR	Pfam: PF00047; I9; 4.	
DR	PRINTS; PR00014; FNTYPEI.II.	
DR	SMART; SM00060; FN3; 6.	
DR	SMART; SM00408; IGC2; 3.	
DR	PROSITE; PS50835; I9-LIKE; 4.	
DR	Cell adhesion; Repeat; Signal; Transmembrane; Immunoglobulin domain;	
KW	Glycoprotein.	
FT	NON_TER	1
FT	POTENTIAL.	1
FT	SIGNAL	2
FT	CHAIN	<1
FT		1377
FT		3

```

      137 LAPPQAKQFLISPPSPVWGQPINADP---VLNYDLLAVAKLGPGEYELHAGTES 193
      | : | : | | | | | | | | | : | : | : | : | : | : |
      85 LQVPKEKNWLISPPGPPVGEPIVEESPNSQHLADQLKDELGNALLNDHSAG--- 141

      194 TPSVVVHVCDSDIEEDDPKTS 216
      | : | : | : | : | : |
      142 -PQIVIS-----EHNNTKETSP 157

```

RESULT 11

```

YAP1_HUMAN STANDARD; PRT; 454 AA.
ID YAP1_HUMAN
AC P46937;
AC 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 65 kDa Yes-associated protein (YAP65).
DE YAP1 OR YAP65.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxId=9606;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Lung;
RC MEDLINE=95301570; Pubmed=7782338;
RC Sudol M., Bork P., Einbond A., Kastury K., Druck T., Negrini M.,
RA Huebner K., Lehman D.;
RT "Characterization of the mammalian YAP (Yes-associated protein) gene
RT and its role in defining a novel protein module, the WW domain.";
RL J. Biol. Chem. 270:14733-14741(1995).
CC [2]
CC INTERACTION WITH WBPI AND WBP2
CC MEDLINE=97347517; Pubmed=9202023;
CC Chen H.I., Einbond A., Kwak S.-J., Linn H., Koepf E., Peterson S.,
CC Kelly J.W., Sudol M.;
CC "Characterization of the WW domain of human Yes-associated protein and
CC its polyproline containing ligands.";
CC RT J. Biol. Chem. 272:17070-17077(1997).
CC RL -1- SUBUNIT: Binds to the SH3 domain of the YES kinase. Binds to WBP1
CC and WBP2.
CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 WW domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X80507; CAA56672.1; -.
CC DR PIR: A56954; A56954.
CC DR PDB: 1JW0; 21-DEC-01.
CC DR PDB: 1K90; 28-DEC-01.
CC DR PDB: 1K9R; 28-DEC-01.
CC DR Genew: HGNC:16262; YAP1.
CC DR GIM: 606608; -.
CC DR GO: 00005515; F:protein binding activity; TAS.
CC DR InterPro: IPR005153; Mbth.
CC DR InterPro: IPR001202; WW_Rsp5_WMP.
CC DR Pfam: PF00397; WW; 1.
CC DR SMART: SM00456; WW; 1.
CC DR PROSITE: PS01159; WW_DOMAIN_1; 1.
CC DR PROSITE: PS50020; WW_DOMAIN_2; 1.
CC DR Phosphorylation; 3D-structure.
CC KW

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 18:58:43 ; Search time 60 Seconds  
(without alignments)  
514.651 Million cell updates/sec

Title: US-10-030-613-1

Perfect score: 1254

Sequence: 1 MRSPPGQGHVPEDEGLFLLC.....SPKPKIIQTRRGLPPSVSN 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	913	72.8	192	10	US-09-782-953-18
2	913	72.8	192	10	US-09-782-953-19
3	905	72.2	197	10	US-09-782-953-9
4	905	72.2	197	10	US-09-782-953-10
5	729.5	58.2	241	10	US-09-782-953-21
6	729.5	58.2	241	10	US-09-782-953-22
7	698.5	55.7	212	10	US-09-782-953-24
8	698.5	55.7	212	10	US-09-782-953-25
9	629	50.2	197	10	US-09-782-953-12
10	629	50.2	197	10	US-09-782-953-13
11	629	50.2	197	10	US-09-782-953-15
12	629	50.2	197	10	US-09-782-953-16
13	623	49.7	198	10	US-09-782-953-6
14	623	49.7	198	10	US-09-782-953-7
15	620.5	49.5	198	10	US-09-782-953-3

16	620.5	49.5	198	10	US-09-782-953-4	Sequence 4, Appli
17	282	22.5	142	9	US-09-925-302-790	Sequence 790, App
18	222	17.7	58	9	US-09-864-761-35379	Sequence 35379, A
19	222	17.7	58	9	US-09-864-761-43076	Sequence 43076, A
20	171	13.6	56	9	US-09-864-761-34111	Sequence 34111, A
21	112	8.9	32	15	US-10-194-155-8	Sequence 8, Appli
22	90	7.2	4019	10	US-09-738-973-425	Sequence 425, App
23	90	7.2	4019	10	US-09-854-133-425	Sequence 425, App
24	90	7.2	4019	10	US-10-144-649A-425	Sequence 425, App
25	84.5	6.7	212	15	US-10-278-173-122	Sequence 122, App
26	84.5	6.7	212	15	US-10-295-403-16	Sequence 16, Appl
27	84	6.7	3063	15	US-10-177-293-61	Sequence 63, Appl
28	83.5	6.7	229	11	US-09-934-455-508	Sequence 508, App
29	83.5	6.7	434	15	US-10-102-806-506	Sequence 506, App
30	83.5	6.7	659	15	US-10-211-962-12	Sequence 12, Appl
31	83.5	6.7	3063	12	US-10-301-822-26	Sequence 26, Appl
32	83.5	6.7	3063	15	US-10-177-293-61	Sequence 61, Appl
33	82	6.5	526	10	US-09-801-368-362	Sequence 362, App
34	81.5	6.5	916	10	US-09-919-497-64	Sequence 64, Appl
35	81.5	6.5	916	10	US-09-925-300-1610	Sequence 1610, Ap
36	80.5	6.4	296	10	US-09-789-054A-10	Sequence 10, Appl
37	80.5	6.4	932	10	US-09-754-997A-2	Sequence 2, Appli
38	80.5	6.4	1252	10	US-09-908-193-21	Sequence 21, Appl
39	80.5	6.4	2701	15	US-10-171-311-83	Sequence 83, Appl
40	80.5	6.3	1706	15	US-10-024-450-2	Sequence 2, Appli
41	79.5	6.3	1706	15	US-10-142-650-3	Sequence 3, Appli
42	79.5	6.3	434	15	US-10-192-634-7	Sequence 7, Appli
43	79	6.3	434	15	US-09-854-549-7	Sequence 7, Appli
44	79	6.3	559	9	US-09-836-561-6	Sequence 6, Appli
45	79	6.3	559	9	US-09-836-561-6	Sequence 6, Appli

## ALIGNMENTS

## RESULT 1

US-09-782-953-18  
; Sequence 18, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; FILE REFERENCE: CALCIUMINURIN INTERACTING PROTEIN (MCIP)  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: US/09782,953  
; PRIOR FILING DATE: 60/216,601  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-953-18

Query Match	72.8%	Score 913	DB 10	Length 192
Best Local Similarity	93.0%	Pred. No. 4.8e-84		
Matches 174	Conservative 3	Mismatches 10	Indels 0	Gaps 0
QY	48	NSLFACNVHQSVEGEESKEPEGLFRYDDCVTFQLEKSFRRVRINFSPKSAARIE	107	
Db	6	STLVACVVDVEFTNQEVKEKFGGLFRYDDCVTFQLEKSFRRVRINFSPKSAARIE	65	
QY	108	LHETFRGKKLYIAQVOTPETDGDKLHLAPPAPAKQFLISPPSPVGVQNDATPV	167	
Db	66	LHETFRGKKLYIAQVOTPETDGDKLHLAPPAPAKQFLISPPSPVGVQNDATPV	125	
QY	168	LNVDLLYAVAKLGRGEKYEHLHAGTESPVSVVHVCDSDIEEDPKTSPKPKIIQTRRP	227	
Db	126	LNVDLLYAVAKLGRGEKYEHLHAGTESPVSVVHVCDSDIEEDPKTSPKPKIIQTRRP	185	

us-10-030-613-1.1.rapb

Tue Aug 26 09:11:02 2003

QY 48 NSLFACNVHQSVEFEGESKEFGLFRTYDDCVTFQLFKSRFRVRINFNPKSAARARIE 107  
 Db 11 STLAVCVVDVEFTNQEVKEFGLFRTYDDCVTFQLFKSRFRVRINFNPKSAARARIE 70  
 QY 108 LHETQFRGKKLKYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVQPIINDATPV 167  
 Db 71 LHETQFRGKKLKYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVQPIINDATPV 130  
 QY 168 LNYDLLYAVAKLGPGEKYEHLHAGTSTPSVVHVHVCDSIDEEEDPKTSKPKEIOTRRPG 227  
 Db 131 LNYDLLYAVAKLGPGEKYEHLHAGTSTPSVVHVHVCDSIDEEEDPKTSKPKEIOTRRPG 190  
 QY 228 LPPSVSN 234  
 Db 191 LPPSVSN 197

RESULT 4

US-09-782-953-10  
 ; Sequence 10, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 ; FILE REFERENCE: UTSD:674P21  
 ; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; CURRENT FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 10  
 ; LENGTH: 197  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-782-953-10

Query Match 72.2%; Score 905; DB 10; Length 197;  
 Best Local Similarity 91.4%; Pred. No. 3.2e-83;  
 Matches 171; Conservative 8; Mismatches 8; Indels 0; Gaps 0;  
 QY 48 NSLFACNVHQSVEFEGESKEFGLFRTYDDCVTFQLFKSRFRVRINFNPKSAARARIE 107  
 Db 11 STLAVCVVDVEFTNQEVKEFGLFRTYDDCVTFQLFKSRFRVRINFNPKSAARARIE 70  
 QY 108 LHETQFRGKKLKYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVQPIINDATPV 167  
 Db 71 LHETQFRGKKLKYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVQPIINDATPV 130  
 QY 168 LNYDLLYAVAKLGPGEKYEHLHAGTSTPSVVHVHVCDSIDEEEDPKTSKPKEIOTRRPG 227  
 Db 131 LNYDLLYAVAKLGPGEKYEHLHAGTSTPSVVHVHVCDSIDEEEDPKTSKPKEIOTRRPG 190  
 QY 228 LPPSVSN 234  
 Db 191 LPPSVSN 197

RESULT 5

US-09-782-953-21  
 ; Sequence 21, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 ; FILE REFERENCE: UTSD:674P21  
 ; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; CURRENT FILING DATE: 2001-02-13

QY 228 LPPSVSN 234  
 Db 186 LPPSVSN 192  
 RESULT 2  
 US-09-782-953-19  
 ; Sequence 19, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 ; FILE REFERENCE: UTSD:674P21  
 ; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; CURRENT FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 19  
 ; LENGTH: 192  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-782-953-19

Query Match 72.8%; Score 913; DB 10; Length 192;  
 Best Local Similarity 93.0%; Pred. No. 4.8e-84;  
 Matches 174; Conservative 3; Mismatches 10; Indels 0; Gaps 0;  
 QY 48 NSLFACNVHQSVEFEGESKEFGLFRTYDDCVTFQLFKSRFRVRINFNPKSAARARIE 107  
 Db 6 STLAVCVVDVEFTNQEVKEFGLFRTYDDCVTFQLFKSRFRVRINFNPKSAARARIE 65  
 QY 108 LHETQFRGKKLKYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVQPIINDATPV 167  
 Db 66 LHETQFRGKKLKYFAQVQTPETDGDGKLHAPPQAKQFLISPPSPVGVQPIINDATPV 125  
 QY 168 LNYDLLYAVAKLGPGEKYEHLHAGTSTPSVVHVHVCDSIDEEEDPKTSKPKEIOTRRPG 227  
 Db 126 LNYDLLYAVAKLGPGEKYEHLHAGTSTPSVVHVHVCDSIDEEEDPKTSKPKEIOTRRPG 185  
 QY 228 LPPSVSN 234  
 Db 186 LPPSVSN 192

RESULT 3

US-09-782-953-9  
 ; Sequence 9, Application US/09782953  
 ; Patent No. US20020150953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLIAMS, R. SANDERS  
 ; APPLICANT: ROTHERMEL, BEVERLY  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
 ; FILE REFERENCE: UTSD:674P21  
 ; CURRENT APPLICATION NUMBER: US/09/782,953  
 ; CURRENT FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/216,601  
 ; PRIOR FILING DATE: 2000-07-07  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 9  
 ; LENGTH: 197  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-782-953-9

Query Match 72.2%; Score 905; DB 10; Length 197;  
 Best Local Similarity 91.4%; Pred. No. 3.2e-83;  
 Matches 171; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

Result No.	Query No.	Query %			DB	ID	Description
		Match	Length	Score			
1	1	913	72.8	197	1	CCP2_HUMAN	Q14206 homo sapien
2	2	905	72.2	197	1	CCP2_MOUSE	Q9jhq2 mus musculus
3	3	739.5	58.5	239	1	CCP3_MOUSE	Q9jkk0 mus musculus
4	4	729.5	58.2	241	1	CCP3_HUMAN	Q9uka8 homo sapien
5	5	632	50.4	197	1	CCP1_CRIGR	Q53847 cricetus
6	6	623	50.2	197	1	CCP1_HUMAN	P53805 homo sapien
7	7	622.5	49.6	198	1	CCP1_MOUSE	Q9jhq6 mus musculus
8	8	432.5	34.5	292	1	SRA_DROME	Q9xz18 drosophila
9	9	351.5	28.0	207	1	CCPL_CAPEL	P53806 caenorhabdi
10	10	107.5	8.6	163	1	YA9F_SCHPO	Q09791 schizosacch
11	11	98	7.8	454	1	YAP1_HUMAN	P46937 homo sapien
12	12	97.5	7.8	1377	1	NEOL_RAT	P07603 rattus norv
13	13	89	7.1	3726	1	ABF1_MOUSE	Q61329 mus musculus
14	14	88	7.0	3329	1	BRC1_MOUSE	P97929 mus musculus
15	15	86.5	6.9	1331	1	MANB_CALSA	P22533 caldocellum
16	16	86	6.9	668	1	PAB5_ARATH	Q35196 arabidopsis
17	17	85	6.8	1232	1	YOO5_CAPEL	P34643 caenorhabdi
18	18	84	6.7	1206	1	PM14_MOUSE	Q05859 mus musculus
19	19	84	6.7	1468	1	FMN1_MOUSE	Q05860 mus musculus
20	20	83.5	6.7	3063	1	FMN1_MOUSE	Q93715 homo sapien
21	21	83	6.6	408	1	CALC_HUMAN	Q35368 mus musculus
22	22	83	6.6	853	1	IF13_MOUSE	P10383 drosophila
23	23	83	6.6	5120	1	OTU_DROME	Q9pu36 gallus gall
24	24	82.5	6.6	457	1	PCLO_CHICK	P20768 suncus muri
25	25	82.5	6.6	663	1	MUC_SUNMU	P27594 sus scrofa
26	26	82	6.5	425	1	XTR1_PIG	P30119 epstein-bar
27	27	82	6.5	434	1	BC13_MOUSE	P59017 mus musculus
28	28	82	6.5	526	1	CAP_YEAST	P17555 saccharomyc
29	29	81.5	6.5	905	1	XPF_HUMAN	Q92889 homo sapien
30	30	81.5	6.5	2319	1	NTG3_RAT	Q9rl72 rattus norv
31	31	81	6.5	639	1	CALC_RABIT	Q28902 orthotolagus
32	32	80.5	6.4	587	1	SYD_SFRCO	Q9f323 streptomyces
33	33	80.5	6.4	611	1	PE34_YEAST	P39684 saccharomyce

Db 11 STLAVCVVDVEVFTNOEVKEKFGGLFTYDDCVTFQLFKSFRRVRINFNSPKSAARARIE 70  
 Qy 108 LHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSSPPVQWQINDATPV 167  
 Db 71 LHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSSPPVQWQINDATPV 130  
 Qy 168 LNYDLLYAVAKLPGGEKYLHAGTSTPSSVVHVHVCDSIDIEEDPKTSPKPKIIQTRRP 227  
 Db 131 LNYDLLYAVAKLPGGEKYLHAGTSTPSSVVHVHVCDSIDIEEDPKTSPKPKIIQTRRP 190  
 Qy 228 LPPSVSN 234  
 Db 191 LPPSVSN 197

## RESULT 2

CCP2\_MOUSE STANDARD; PRT; 197 AA.  
 AC Q9JHG2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Calcipressin 2 (Down syndrome candidate region 1-like protein 1)  
 DE (Myocyte-enriched calcineurin interacting protein 2) (MCIP2).  
 GN DSCR1L1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20187590; PubMed=10722714;  
 RA Rothermel B., Vega R.B., Yang J., Wu H., Bassel-Duby R.,  
 RA Williams R.S.;  
 RT "A protein encoded within the Down syndrome critical region is  
 RT enriched in striated muscles and inhibits calcineurin signaling.";  
 RL J. Biol. Chem. 275:8719-8725(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=BALB/C; TISSUE=Brain;  
 RX MEDLINE=20534792; PubMed=11080588;  
 RA Striappoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;  
 RA "The murine DSCR1-like (Down syndrome candidate region 1) gene family:  
 RT conserved synteny with the human orthologous genes.";  
 RL Gene 257:223-232(2000).  
 CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses  
 CC by binding to the catalytic domain of calcineurin A. Could play a  
 CC role during central nervous system development.  
 CC -!- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle  
 CC and brain. Lower expression in all other tissues.  
 CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF237791; AAF63487.1;  
 CC EMBL; AF237887; AAF62538.1;  
 CC MGD; MGI:1858219; Dscr1l1.  
 CC InterPro: IPR006931; Calcipressin.  
 CC Pfam: PF04847; Calcipressin; 1.  
 SQ SEQUENCE 197 AA; 22025 MW; CE1306B7B03E70F1 CRC64;

Query Match 72.2%; Score 905; DB 1; Length 197;  
 Best Local Similarity 91.4%; Pred. No. 1e-70; Indels 0; Gaps 0;  
 Matches 171; Conservative 8; Mismatches 8;  
 Qy 48 NSLFCNVHOSVFGEESKEKFGGLFTYDDCVTFQLFKSFRRVRINFNSPKSAARARIE 107

Db 11 STLAVCVVDVEVFTNOEVKEKFGGLFTYDDCVTFQLFKSFRRVRINFNSPKSAARARIE 70  
 Qy 108 LHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSSPPVQWQINDATPV 167  
 Db 71 LHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSSPPVQWQINDATPV 130  
 Qy 168 LNYDLLYAVAKLPGGEKYLHAGTSTPSSVVHVHVCDSIDIEEDPKTSPKPKIIQTRRP 227  
 Db 131 LNYDLLYAVAKLPGGEKYLHAGTSTPSSVVHVHVCDSIDIEEDPKTSPKPKIIQTRRP 190  
 Qy 228 LPPSVSN 234  
 Db 191 LPPSVSN 197

## RESULT 3

CCP3\_MOUSE STANDARD; PRT; 239 AA.  
 AC Q9JUK0; Q9CX87;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Calcipressin 3 (Down syndrome candidate region 1-like protein 2)  
 DE (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).  
 GN DSCR1L2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=BALB/C; TISSUE=Brain;  
 RA Striappoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;  
 RT "The murine DSCR1-like gene family.";  
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 40-239 FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=Embryonic head;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,  
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,  
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses  
 CC by binding to the catalytic domain of calcineurin A. Could play a  
 CC role during central nervous system development (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF237888; AAF62539.1;  
 CC

DR EMBL; AK019377; BAB31587.1; -  
 DR MGD; MGI:1858220; Dscril2.  
 DR InterPro; IPR006931; Calcipressin.  
 DR Pfam; PF04847; Calcipressin; 1.  
 SQ SEQUENCE 239 AA; 27153 MW; 1B2687B47BB4D272 CRC64;

Query Match 58.5%; Score 733.5; DB 1; Length 239;  
 Best Local Similarity 67.1%; Pred. No. 6.6e-56;  
 Matches 143; Conservative 24; Mismatches 43; Indels 3; Gaps 2;

QY 19 LCICDRDWAATRCFA--EEAFQAITDFNDLNSFACNVHQSVEGESEKKEGFLRTY 76  
 DB 16 LCSSDQEEEMVFNEDGEGLEMDLSDPTSLFACSVHEAFVEQEKERFEALFTLY 75  
 QY 77 DCVTFQLFKSFRRVRINFSPKSAARARIELHETQPRGKKLYFAQVQTPETDGDKLH 136  
 DB 76 DDVTFQLFKSFRRVRINFSPKSAARARIELHETQPRGKKLYFAQVQTPETDGDKLH 136  
 QY 137 LAPPOPAKOFILSPSPSPVGVQPIQNDATPVNLDLYAVAKLGPGEKYLHAGTSTPS 196  
 DB 136 LLPQPTKQFLISPPSPVGVQKQSEDAMPVINDLLCAVSKLGPGEKYLHAGTSTPS 195  
 QY 197 VVHVCDSDIEEDDKTSPKPKIIQTRRPGLP 229  
 DB 196 VVHVCESETEEDTK-NPKQKITOTRRPEAP 227

RESULT 4  
 CCP3\_HUMAN  
 ID CCP3\_HUMAN STANDARD; PRT; 241 AA.  
 AC Q9UKA8; Q9UKA7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Calcipressin 3 (Down syndrome candidate region 1-like protein 2)  
 DE (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).  
 GN DSCRIL2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE-Peripheral blood, and placenta;  
 RX MEDLINE=2021370; PubMed=10756093;  
 RA Strippoli P., Lenzi L., Petrini M., Carinci P., Zannotti M.;  
 RT and ZAKI-4: characterization from yeast to human and identification of  
 RL DSCR1-like 2, a novel human member (DSCRIL2).";  
 RN Genomics 64:252-263(2000).  
 RP SEQUENCE OF 66-241 FROM N.A.  
 RA Bagguley C.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses  
 CC by binding to the catalytic domain of calcineurin A. Could play a  
 CC role during central nervous system development (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q9UKA8-1; Sequence-Displayed;  
 CC Name=2;  
 CC IsoId=Q9UKA8-2; Sequence=VSP\_001319;  
 CC -!- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle  
 CC kidney, liver and peripheral blood leukocytes. Lower expression in  
 CC all other tissues.  
 CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/  
 CC or send an email to license@isb-sib.ch).

CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; AF176116; AAF01684.1; -  
 CC EMBL; AF176117; AAF01685.1; -  
 CC EMBL; AL034582; CNB72298.1; -  
 CC Genew; HGNC:3042; DSCRIL2.  
 CC MIM; 605860; -  
 CC GO; GO:0003723; F:RNA binding activity; TAS.  
 CC GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.  
 CC InterPro; IPR006931; Calcipressin.  
 CC Pfam; PF04847; Calcipressin; 1.  
 KW Alternative splicing.  
 FT VARSPLIC 124 133 Missing (in isoform 2).  
 FT -----  
 SQ SEQUENCE 241 AA; 27492 MW; 91AB619F5E443FBD CRC64;  
 Query Match 58.2%; Score 729.5; DB 1; Length 241;  
 Best Local Similarity 66.1%; Pred. No. 1.5e-55;  
 Matches 144; Conservative 22; Mismatches 49; Indels 3; Gaps 2;

QY 19 LCICDRDWAATRCFA--EEAFQAITDFNDLNSFACNVHQSVEGESEKKEGFLRTY 76  
 DB 16 LCSTDQEEEMIFGENEDDLDMDLSDPTSLFACSVHEAFVEQEKERFEALFTY 75  
 QY 77 DCVTFQLFKSFRRVRINFSPKSAARARIELHETQPRGKKLYFAQVQTPETDGDKLH 136  
 DB 76 DDVTFQLFKSFRRVRINFSPKSAARARIELHETQPRGKKLYFAQVQTPETDGDKLH 136  
 QY 137 LAPPOPAKOFILSPSPSPVGVQPIQNDATPVNLDLYAVAKLGPGEKYLHAGTSTPS 196  
 DB 136 LLPQPTKQFLISPPSPVGVQKQSEDAMPVINDLLCAVSKLGPGEKYLHAGTSTPS 195  
 QY 197 VVHVCDSDIEEDDKTSPKPKIIQTRRPGLP 234  
 DB 196 VVHVCESETEEDTK-NPKQKITOTRRPEAP 232

RESULT 5  
 CCP1\_CRIGR  
 ID CCP1\_CRIGR STANDARD; PRT; 197 AA.  
 AC O35847;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Calcipressin 1 (Down syndrome critical region protein 1 homolog)  
 DE (Oxidative-induced protein Adapt78).  
 GN DSCR1 OR ADAPT78.  
 OS Cricetus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 CC Cricetus.  
 CC NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=97329095; PubMed=9185608;  
 RA Crawford D.R., Leahy K.P., Abramova N., Lan L., Wang Y., Davies K.J.;  
 RT "Hamster adapt78 mRNA is a Down syndrome critical region homologue  
 RT that is inducible by oxidative stress.";  
 RL Arch. Biochem. Biophys. 342:6-12(1997).  
 CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses  
 CC by binding to the catalytic domain of calcineurin A. Could play a  
 CC role during central nervous system development (By similarity).  
 CC -!- INDUCTION: OXIDANT-INDUCIBLE.  
 CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/  
 CC or send an email to license@isb-sib.ch).

Tue Aug 26 09:11:04 2003

```

CC -----
DR EMBL: U60263; AAB68517.1; -
DR InterPro: IP0006931; Calcipressin.
DR Pfam: PF04847; Calcipressin; 1.
DR SQ SEQUENCE 197 AA; 22674 MW; 0EFC68FA55826439 CRC64;

Query Match 50.4%; Score 632; DB 1; Length 197;
Best Local Similarity 63.7%; Pred. NO 2.6e-47;
Matches 123; Conservative 24; Mismatches 38; Indels 8; Gaps 2;

QY 42 DNDLPNSLPACNVHVSFEVGESEKFEGLFRTYDCTFQLFKSFRRVRINFSNPKSA 101
DB 5 DFNFSLLACVANGDVFSESETRAKFSLFTYDKDITTFYFKSFKVRINFSNPLSA 64
QY 102 ARARIEHETQFRGKLLKLYFAQVQTPETDGLHLAPPQAKOFLISPPSPVGVQWPI 161
DB 102 ARARIEHETQFRGKLLKLYFAQVQTPETDGLHLAPPQAKOFLISPPSPVGVQWPI 161
QY 65 ADARLQHLKTEFLGKEMKLYFAQTLHGSS-----HLAPPNPDKOFLISPPSPVGVQW 120
DB 65 ADARLQHLKTEFLGKEMKLYFAQTLHGSS-----HLAPPNPDKOFLISPPSPVGVQW 120
QY 162 NDATPVLNVDLYAVAKLPGCKEYELHAGTSTSVVHVCDSDIEEDPKTS-----PK 217
DB 162 NDATPVLNVDLYAVAKLPGCKEYELHAGTSTSVVHVCDSDIEEDPKTS-----PK 217
QY 121 EDATPVINYDLYALSKLPGCKEYELHAATDTTSPVVHVCDSDIEEDPKTS-----PK 180
DB 121 EDATPVINYDLYALSKLPGCKEYELHAATDTTSPVVHVCDSDIEEDPKTS-----PK 180
QY 218 PKIOTRRRGLPP 230
DB 181 PKIOTRRRPEYTP 193

RESULT 6
CCPI_HUMAN STANDARD: PRT; 197 AA.
ID CCPI_HUMAN STANDARD: Q96R03; Q9B069; Q9UF15; Q9OME4;
AC P53805; O00582; O00583; Q96R03; Q9B069; Q9UF15; Q9OME4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Calcipressin 1 (Down syndrome critical region protein 1) (Myocyte-
DE enriched calcineurin interacting protein 1) (MCIP1) (Adapt78).
DE DSCR1 OR DSC1 OR ADAPT78.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RX MEDLINE=96121593; PubMed=8595418;
RA Fuentes J.-J., Pritchard M.A., Planas A.M., Bosch A., Ferrer I.,
RA Estivill X.;
RT "A new human gene from the Down syndrome critical region encodes a
RT proline-rich protein highly expressed in fetal brain and heart.";
RL Hum. Mol. Genet. 4:1935-1944(1995).
[2]
RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=97468152; PubMed=9325050;
RA Fuentes J.-J., Pritchard M.A., Estivill X.;
RA "Genomic organization, alternative splicing, and expression patterns
RA of the DSCR1 (Down syndrome candidate region 1) gene.";
RT Genomics 44:358-361(1997).
[3]
RN SEQUENCE FROM N.A. (ISOFORM 3).
RA Hua F., Wu J., Zhou Y., Zhang B., Peng X., Qiang B., Yuan J.,
RA Qiang B.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Ioshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

```

```

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling R., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Small D.E.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length
RA human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [5]
RN SEQUENCE OF 1-164 FROM N.A. (ISOFORM 2).
RC TISSUE=Breast; Leahy K.P., Davies K.J.A.;
RA Crawford D.R., a calcium and oxidant-inducible RNA.";
RT "Adapt78, a calcium and oxidant-inducible RNA.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
[6]
RN CHARACTERIZATION.
RX MEDLINE=20320698; PubMed=10861295;
RA Fuentes J.J., Genesca L., Kingsbury T.J., Cunningham K.W.,
RA Perez-Riba M., Estivill X., de la Luna S.;
RT "DSCR1, overexpressed in Down syndrome, is an inhibitor of
RT calcineurin-mediated signaling pathways.";
RL Hum. Mol. Genet. 9:1681-1690(2000).
CC "-!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1;
CC IsoId=P53805-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P53805-2; Sequence=VSP_001314;
CC Name=3;
CC IsoId=P53805-3; Sequence=VSP_001315;
CC Name=4;
CC IsoId=P53805-4; Sequence=VSP_001316;
CC "-!- TISSUE SPECIFICITY: Highly expressed heart, brain and skeletal
CC muscle. Also expressed in all other tissues.
CC -!- INDUCTION: By calcium.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U28833; AAB81557.1; -
CC EMBL: U85265; AAB84370.1; -
CC EMBL: U85266; AAB84371.2; -
CC EMBL: U85267; AAB84372.1; -
CC EMBL: AF400429; AAK92478.1; -
CC EMBL: BC002864; AAH02864.1; -
CC EMBL: U53821; AAF21218.1; -
CC Genew; HGNC:3040; DSCR1.
CC MIM; 602917; -
CC GO; GO:0005634; C:nucleus; TAS.
CC GO; GO:0003700; F:transcription factor activity; TAS.
CC GO; GO:0007417; P:central nervous system development; TAS.
CC GO; GO:0008015; P:circulation; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro; IP0006931; Calcipressin.
CC Pfam; PF04847; Calcipressin.
CC Alternative splicing.
CC DOMAIN 166 174
CC VARSPLIC 1 28
CC POLY-GLU.
CC MERSVDQLQPSATIAChLPRFVDCGLC -> MHERNENYS
CC FSSLIACVANSIDFSESET (in isoform 2).
CC /FTId=VSP_001314.

```







[illegible]

	Query Match	55.7%	Score 698.5;	DB 10;	Length 212;
	Best Local Similarity	67.7%;	Pred. No. 2.3e-62;		
	Matches 136; Conservative	20; Mismatches 34;	Indels 11; Gaps 2;		
QY	34	EAAFAQITDNDLPNSLFACNVHQSVPFEGESKEKFGLEFTDYDDCVTFOLFESFRVRI	93		
	:   :	:        :    :	:		
Dd	14	EDDLDEMDLSDUDTSLFACSVAHEAFPEAKRERFEALFTIYDDQVTFOLEKSFRVRI	73		
	:	:        :    :	:		

QY 94 NFNPKSAARARIELHETQFGKGLKLYFAOVQTPETDGDKLHLAPPQAKQFLISPPSS 153  
DB 74 NFNPKSAARARIELHETQFGKGLKLYFAOVQTPETDGDKLHLAPPQAKQFLISPPAS 123  
QY 154 PPGVQNPINDATPVNLDLLYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDIEEDDPK 213  
DB 124 PPGVQKQSEDAMPVINYDLLCAVSKLGPGEKYLHAGTSTPSVVVHVCDSDIEEDDPK 183  
QY 214 TSPKPKIOTRRPGLPPSVSN 234  
DB 184 -NPKQKIAQTRRRPDPPTAALN 203

RESULT 9  
US-09-782-953-12  
; Sequence 12, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-953-12

Query Match 50.2%; Score 629; DB 10; Length 197;  
Best Local Similarity 62.7%; Pred. No. 2.1e-55;  
Matches 121; Conservative 27; Mismatches 37; Indels 8; Gaps 2;

QY 42 DFNDLPNSLFACNVHQSVEGEESKEFGLFTYDDCVTFQFLKSFRRVRINFSPKSA 101  
DB 5 NFNYSFSLIACVANSDFSESETRAKFESLFTYDKDITTFQYFKSKVRINFSPKSA 64  
QY 102 ARARIELHETQFGKGLKLYFAOVQTPETDGDKLHLAPPQAKQFLISPPSPVGVQPI 161  
DB 65 ADARLQLHKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPSPVGVQKQV 120  
QY 162 NDATPVNLDLLYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDIEEDDPKTS----PK 217  
DB 121 EDATPVINYDLYAISKLGPGKYLHAGTSTPSVVVHVCDSDIEEDDPKTS----PK 180  
QY 218 PKIOTRRPGLPP 230  
DB 181 PKIOTRRPEYTP 193

RESULT 11  
US-09-782-953-15  
; Sequence 15, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-953-15

Query Match 50.2%; Score 629; DB 10; Length 197;  
Best Local Similarity 62.7%; Pred. No. 2.1e-55;  
Matches 121; Conservative 26; Mismatches 38; Indels 8; Gaps 2;

QY 42 DFNDLPNSLFACNVHQSVEGEESKEFGLFTYDDCVTFQFLKSFRRVRINFSPKSA 101  
DB 5 DLODLESATLACHLDPRVFDGLCRAKFESLFTYDKDITTFQYFKSKVRINFSPKSA 64  
QY 102 ARARIELHETQFGKGLKLYFAOVQTPETDGDKLHLAPPQAKQFLISPPSPVGVQPI 161  
DB 65 ADARLQLHKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPSPVGVQKQV 120  
QY 162 NDATPVNLDLLYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDIEEDDPKTS----PK 217  
DB 121 EDATPVINYDLYAISKLGPGKYLHAGTSTPSVVVHVCDSDIEEDDPKTS----PK 180  
QY 218 PKIOTRRPGLPP 230  
DB 181 PKIOTRRPEYTP 193

RESULT 12  
US-09-782-953-16  
; Sequence 16, Application US/09782953

QY 94 NFNPKSAARARIELHETQFGKGLKLYFAOVQTPETDGDKLHLAPPQAKQFLISPPSS 153  
DB 74 NFNPKSAARARIELHETQFGKGLKLYFAOVQTPETDGDKLHLAPPQAKQFLISPPAS 123  
QY 154 PPGVQNPINDATPVNLDLLYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDIEEDDPK 213  
DB 124 PPGVQKQSEDAMPVINYDLLCAVSKLGPGEKYLHAGTSTPSVVVHVCDSDIEEDDPK 183  
QY 214 TSPKPKIOTRRPGLPPSVSN 234  
DB 184 -NPKQKIAQTRRRPDPPTAALN 203

RESULT 9  
US-09-782-953-12  
; Sequence 12, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-953-12

Query Match 50.2%; Score 629; DB 10; Length 197;  
Best Local Similarity 62.7%; Pred. No. 2.1e-55;  
Matches 121; Conservative 27; Mismatches 37; Indels 8; Gaps 2;

QY 42 DFNDLPNSLFACNVHQSVEGEESKEFGLFTYDDCVTFQFLKSFRRVRINFSPKSA 101  
DB 5 NFNYSFSLIACVANSDFSESETRAKFESLFTYDKDITTFQYFKSKVRINFSPKSA 64  
QY 102 ARARIELHETQFGKGLKLYFAOVQTPETDGDKLHLAPPQAKQFLISPPSPVGVQPI 161  
DB 65 ADARLQLHKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPSPVGVQKQV 120  
QY 162 NDATPVNLDLLYAVAKLGPGEKYLHAGTSTPSVVVHVCDSDIEEDDPKTS----PK 217  
DB 121 EDATPVINYDLYAISKLGPGKYLHAGTSTPSVVVHVCDSDIEEDDPKTS----PK 180  
QY 218 PKIOTRRPGLPP 230  
DB 181 PKIOTRRPEYTP 193

RESULT 10  
US-09-782-953-13  
; Sequence 13, Application US/09782953  
; Patent No. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE  
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)  
; FILE REFERENCE: UTSD:674P21  
; CURRENT APPLICATION NUMBER: US/09/782,953  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 60/216,601  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1

```

; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-16

Query Match          50.2%; Score 629; DB 10; Length 197;
Best Local Similarity 62.7%; Pred No. 2.1e-55;
Matches 121; Conservative 26; Mismatches 38; Indels 8; Gaps 3

QY 42 DFNDLPSLFCACNVHQSVEGESEKFEGLFRTYDDCVTFQLFKSFRVRINFNSPKSA 107
Db 5 DLQDPSATTACHLDRVFVDGLGCRAKFESLFTYDKDITTFQYFKSFRVRINFNSPESA 64
QY 102 ARARIELHETQFRGKKLKLKYFAQVQPTDGDGKLHLAPPQAKQFLISPPSPVGMQPI 167
Db 65 ADARQLHKTFFLGKEMKLYFAQTLHGSS----HLAPPNDPKQFLISPPASPPVGMQV 120
QY 162 NDATPVNLDLLYAVAKLGGKYEKLHAGTESTPSVWVHVCDSDEEEDPKTS----PK 217
Db 121 EDATPVNLDLLYAVAKLGGKYEKLHAGTDTTPSVWVHVCSDOEKEEEMERMRRPK 180
QY 218 PKIIQTRRRGLPP 230
Db 181 PKIIQTRRRPYTP 193

RESULT 13
US-09-782-953-6
; Sequence 6, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-6

Query Match          49.7%; Score 623; DB 10; Length 198;
Best Local Similarity 62.4%; Pred. No. 8.4e-55;
Matches 123; Conservative 21; Mismatches 41; Indels 12; Gaps 3

QY 42 DFNDLP---NSLFCACNVHQSVEGESEKFEGLFRTYDDCVTFQLFKSFRVRINFNSP 98
Db 2 DFRDPSYNSSLIACVANDVDFSESETRAKFESLFTYDKDTTFQYFKSFRVRINFNSP 61
QY 99 KSAARARIELHETQFRGKKLKLKYFAQVQPTDGDGKLHLAPPQAKQFLISPPSPVGM 158
Db 62 LSAADARLRLHKTEFLGKEMKLYFAOTLHGSS----HLAPPNDPKQFLISPPASPPVGM 117

```

Tue Aug 26 09:11:02 2003

US-09-782-953-3

Query Match	49.5%	Score	620.5	DB	10	Length	198
Best Local Similarity	62.4%	Pred. No.	1.5e-54				
Matches	121	Conservative	22	Mismatches	42	Indels	9
						Gaps	2

QY	42	DENDLPNSLFACNVHQSVFEGESKEKEFGLFRITYDDCVTFQLPKSFRRVRINFSPKSA	101
Db	5	DIQDLPSATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYKSKRRVRINFSPLSA	64
QY	102	ARARIELHETQPRGKLLKLYFAQVQTPETDGDKHLAPPQPAKQFLISPPSSPPVGVQPI	161
Db	65	ADARLRLHKTFLGKEMKLYFAQTLHIGSS---HLAPPNDKQFLISPPASPPVGVQKQV	120
QY	162	NDATPVNLYDLYAVAKLGPGEKYLHAGTESTPSVVHVHVCDSIEEEDPK-----TSP	216
Db	121	EDATPVNLYDLYAISKLGPEKYLHAAATDPTPSVVHVHVCESDQENEEEMERMRKP	180
QY	217	KPKIOTRRRGLPP	230
Db	181	KPKIOTRRPEYTP	194

Search completed: August 22, 2003, 19:08:07  
Job time : 61 secs

FT	DOMAIN	3	1074	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1075	1095	POTENTIAL.
FT	DOMAIN	1096	1377	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	21	114	IG-LIKE C2-TYPE 1.
FT	DOMAIN	121	206	IG-LIKE C2-TYPE 2.
FT	DOMAIN	198	305	IG-LIKE C2-TYPE 3.
FT	DOMAIN	310	395	IG-LIKE C2-TYPE 4.
FT	DOMAIN	405	502	FIBONECTIN TYPE-III 1.
FT	DOMAIN	505	598	FIBONECTIN TYPE-III 2.
FT	DOMAIN	599	698	FIBONECTIN TYPE-III 3.
FT	DOMAIN	704	798	FIBONECTIN TYPE-III 4.
FT	DOMAIN	819	919	FIBONECTIN TYPE-III 5.
FT	DOMAIN	920	1021	FIBONECTIN TYPE-III 6.
FT	DOMAIN	1087	1090	POLY-VAL.
FT	DISULFID	43	98	BY SIMILARITY.
FT	DISULFID	142	190	BY SIMILARITY.
FT	DISULFID	239	289	BY SIMILARITY.
FT	DISULFID	331	379	BY SIMILARITY.
FT	CARBOHYD	42	42	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	179	179	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	295	295	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	439	439	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	458	458	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	608	608	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	684	684	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	878	878	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1377	AA; 150637 MW; E514ED8ABD1A63A9	CRC64;

Query Match		7.8%;	Score 97.5;	DB 1;	Length 1377;
Best Local Similarity		23.7%;	Pred. No. 2.2;		
Matches 49;		Conservative	29;	Mismatches 82;	Indels 47; Gaps 12;

Qy	42	DFNDLPNSLFACNVHQ--SVFGEESKEKEFEG	-----FRTYDDCVTFQLPKSFR	89
Db	1165	DITPVDNSMDS-NIHQRNSYRGHESDSNSTLAGRGRMPKMMFPDSQPPQ	-----FRTYDDCVTFQLPKSFR	89
Qy	90	RVRINFSNPKASARARIELHETQFRKKLYLFAQVQTPETDGDKL	---HLAPPQAPQAF	146
Db	1221	NTPSTDTPMPASSQTCCTDRQDP-EGATSSSYLASSQ-EEDSGQSILPTAHVRSHPCLKSF	1278	
Qy	147	LIS--PPSSPPVGWQIPNDATPVNLDLLYAVAKLPGKEKYLHA	-----GYESTP	195
Db	1279	AVPAIPPGPGPI-YDPALPSTPLLSQAL	-----NHLHLSVKTASIGTLGRSRPP	1327
Qy	196	SVVVHVCDSDIEE-----EDPKTSPKP	218	
Db	1328	MPVVVSAPEVQATRMLEDSSESYEP	1354	

```

RESULT 13
ABFL_MOUSE
ID ABFL_MOUSE STANDARD; PRT; 3726 AA.
AC Q61329;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alpha-fetoprotein enhancer binding protein (AT motif-binding factor)
DE (AR-binding transcription factor 1).
DE ATBFL.
GN Mus musculus (Mouse).
OS
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC NCBI; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NMBL_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/MK X ICR; TISSUE=Brain;
RX MEDLINE=96194902; PubMed=8654949;
RA Ido A., Miura Y., Watanabe M., Sakai M., Inoue Y., Miki T.,
RT Hashimoto T., Morinaga T., Nishi S., Tamaoki T.;
RA "Cloning of the cDNA encoding the mouse ATBFL transcription factor.";
RL Gene 168:227-231(1996).
CC -!- FUNCTION: Transcriptional activator that binds to the AT-rich core
CC sequence of the enhancer element of the AFP gene.

```

```
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 4 homeobox domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D26046; BAA05046.1; -.
DR HSSP; P20263; LOCP.
DR TRANSFAC; T03881; -.
DR MGD; MGI-99948; Atbf1.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR GO; GO:0030182; F:neuron differentiation; TAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00046; homeobox_4.
DR Pfam; PF00096; zf-C2H2; 18.
DR ProDom; PD000010; Homeobox; 4.
DR PROSITE; PS00027; HOMEBOX_1; 2.
DR PROSITE; PS00071; HOMEBOX_2; 4.
DR PROSITE; PS00081; Zinc_Finger_C2H2_1; 15.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_2; 9.
KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Homeobox; Nuclear protein; Repeat.
FT ZN_FING 79 103
FT ZN_FING 282 305 C2H2-TYPE.
FT ZN_FING 641 664 C2H2-TYPE.
FT ZN_FING 672 695 C2H2-TYPE.
FT ZN_FING 727 751 C2H2-TYPE.
FT ZN_FING 805 829 C2H2-TYPE (ATYPICAL).
FT ZN_FING 946 969 C2H2-TYPE (DEGENERATE).
FT ZN_FING 985 1009 C2H2-TYPE (ATYPICAL).
FT ZN_FING 1041 1065 C2H2-TYPE (ATYPICAL).
FT ZN_FING 1089 1113 C2H2-TYPE (ATYPICAL).
FT ZN_FING 1233 1256 C2H2-TYPE (ATYPICAL).
FT ZN_FING 1262 1285 C2H2-TYPE (ATYPICAL).
FT ZN_FING 1370 1395 C2H2-TYPE.
FT ZN_FING 1411 1433 C2H2-TYPE.
FT ZN_FING 1439 1462 C2H2-TYPE.
FT ZN_FING 1555 1579 C2H2-TYPE.
FT ZN_FING 1606 1630 C2H2-TYPE.
FT ZN_FING 1990 2013 C2H2-TYPE.
FT DNA_BIND 2152 2211 HOMEBOX 1.
FT DNA_BIND 2249 2308 HOMEBOX 2.
FT ZN_FING 2335 2358 C2H2-TYPE (ATYPICAL).
FT ZN_FING 2539 2561 C2H2-TYPE.
FT DNA_BIND 2650 2709 HOMEBOX 3.
FT ZN_FING 2720 2743 C2H2-TYPE.
FT DNA_BIND 2952 3011 HOMEBOX 4.
FT ZN_FING 3032 3056 C2H2-TYPE.
FT ZN_FING 3552 3576 C2H2-TYPE.
FT DOMAIN 461 491 POLY-GLU.
FT DOMAIN 771 785 POLY-ALA.
FT DOMAIN 1314 1317 POLY-ALA.
FT DOMAIN 1734 1748 POLY-GLN.
FT DOMAIN 1794 1799 POLY-GLN.
FT DOMAIN 1856 1863 POLY-GLN.
FT DOMAIN 2044 2059 POLY-PRO.
FT DOMAIN 2405 2408 POLY-ALA.
FT DOMAIN 3216 3220 POLY-PRO.
FT DOMAIN 3380 3409 POLY-GLN.
FT DOMAIN 3412 3420 POLY-GLN.
FT DOMAIN 3534 3550 POLY-GLY.
FT DOMAIN 3620 3623 POLY-PRO.
FT DOMAIN 3659 3662 POLY-SER.
SQ SEQUENCE 3726 AA; 406567 MW; 915ACBF58BA727C98 CD064.
```

Query Match		7.1%; Score 89; DB 1; Length 3726;
Best Local Similarity		28.5%; Pred. No. 41;
Matches		43; Conservative 17; Mismatches 65; Indels 26; Gaps 7;
QY	29	TRCFABE-AFOAITDFNDLPSLFACNVHQSVEEGESKEKFEGLFRYDDCVTFQFLPKS 87
DB	1951	TKALLENFGEFVIOYNE--NKQAKKNGKAEQGESLEKLE-----CDSC-----CKL 1998
QY	88	FRVRVNFNSPKSAAR-----ARIELHETQFRGKKLYFAQVOTPE-----TDGDKL 135
DB	1999	FSNILLKSHQEHVQNYFPFQLEKREFAKQYREHYDKLYPLRPQPEPPPPPPPPPL 2058
QY	136	HIAPPQAKOFLISPPSPVQWQINDATP 166
DB	2059	PTAPPQAPASAPAT-PASAPPITSPITAPAP 2088
RESULT 14		
BRC2_MOUSE		
ID	BRC2_MOUSE	STANDARD; PRT; 3329 AA.
AC	P97929; O35922; P97383;	
DT	15-DEC-1998 (Rel. 37, Created)	
DT	15-DEC-1998 (Rel. 37, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	Breast cancer type 2 susceptibility protein.	
GN	BRC2.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
OC	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=129; SV;	
RC	MEDLINE=97217789; PubMed=9063750;	
RA	Connor F., Smith A., Wooster R., Stratton M., Dixon A., Campbell E.,	
RA	Tait T.M., Freeman T., Ashworth A.;	
RT	"Cloning, chromosomal mapping and expression pattern of the mouse	
RT	Brc2 gene.";	
RL	Hum. Mol. Genet. 6:291-300(1997).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=C57BL/6;	
RC	MEDLINE=97237041; PubMed=9119389;	
RA	Sharan S.K., Bradley A.;	
RA	"Murine Brc2: sequence, map position, and expression pattern.";	
RT	Genomics 40:234-241(1997).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=129/SV;	
RC	MEDLINE=97384941; PubMed=9242436;	
RA	McAllister K.A., Haugen-Strano A., Hagevik S., Brownlee H.A.,	
RA	Collins N.K., Futreal P.A., Bennett L.M., Wiseman R.W.;	
RT	"Characterization of the rat and mouse homologues of the BRC2 breast	
RT	cancer susceptibility gene.";	
RL	Cancer Res. 57:3121-3125(1997).	
RN	[4]	
RP	SEQUENCE OF 18-200 FROM N.A.	
RC	MEDLINE=97075121; PubMed=9917547;	
RA	Rajan J.V., Wang M., Marquis S.T., Chodosh L.A.;	
RT	"Brc2 is coordinately regulated with Brcal during proliferation and	
RT	differentiation in mammary epithelial cells.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 93:13078-13083(1996).	
RN	[5]	
RP	SEQUENCE OF 569-625 FROM N.A.	
RC	MEDLINE=97341126; PubMed=9196008;	
RA	McAllister K.A., Ramachandran S., Haugen-Strano A., Fiedorek F.T. Jr.,	
RA	Wiseman R.W.;	
RT	"Genetic mapping of the Brc2 breast cancer susceptibility gene on	
RT	mouse chromosome 5.";	
RL	Mamm. Genome 8:540-541(1997).	
CC	-1- FUNCTION: MAY PARTICIPATE IN A PATHWAY ASSOCIATED WITH THE	
CC	ACTIVATION OF DOUBLE-STRAND BREAK REPAIR AND/OR HOMOLOGOUS	
CC	RECOMBINATION (BY SIMILARITY).	

CC	-1- SUBUNIT: INTERACTS WITH RAD51 (BY SIMILARITY).	
CC	-1- TISSUE SPECIFICITY: WIDELY EXPRESSED. HIGHEST EXPRESSION IN	
CC	CEREBELLUM, TESTIS, ILEUM, APPENDIX, EPIDIDYMIS, OVARY AND MAMMARY	
CC	GLAND. NO EXPRESSION IN LONG.	
CC	-1- DEVELOPMENTAL STAGE: IN THE MAMMARY GLAND, EXPRESSION INCREASES	
CC	DRAMATICALLY DURING PREGNANCY.	
CC	-1- SIMILARITY: Contains 7 BRC2 repeats.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	EMBL; U82270; AAB48306.1; -	
DR	EMBL; U72947; AAB40720.1; -	
DR	EMBL; U65594; AAC23702.1; -	
DR	EMBL; U89652; AAB71377.1; -	
DR	EMBL; U89503; AAC53276.1; -	
DR	PIR; T30904; T30904	
DR	PDB; 1MIU; 25-SEP-02.	
DR	MGI:109337; Brc2.	
DR	GO; GO:0005737; C:cytoplasm; IDA.	
DR	GO; GO:0005634; C:nucleus; ISS.	
DR	GO; GO:0005716; C:synaptonemal complex; ISS.	
DR	GO; GO:0005515; F:protein binding activity; ISS.	
DR	GO; GO:0003697; F:single-stranded DNA binding activity; ISS.	
DR	GO; GO:0003653; F:transcriptional activator activity; ISS.	
DR	GO; GO:0006915; P:apoptosis; ISS.	
DR	GO; GO:0006338; P:chromatin modeling; ISS.	
DR	GO; GO:0006260; P:DNA replication; ISS.	
DR	GO; GO:0000724; P:double-strand break repair via homologous r. . .; ISS.	
DR	GO; GO:0006325; P:establishment and/or maintenance of chromatin. . .; ISS.	
DR	GO; GO:0007093; P:mitotic checkpoint; ISS.	
DR	GO; GO:0007090; P:regulation of S phase of mitotic cell cycle; ISS.	
DR	GO; GO:0045449; P:regulation of transcription; ISS.	
DR	InterPro; IPR002093; BRC2_repeat.	
DR	Pfam; PF00634; BRC2; 7.	
DR	PROSITE; PS0138; BRC2_REPEAT; 6.	
KW	POLYMERISM: Repeat; 3D-structure.	
FT	REPEAT 981 1015	
FT	REPEAT 1192 1226	
FT	REPEAT 1394 1428	
FT	REPEAT 1491 1525	
FT	REPEAT 1623 1657	
FT	REPEAT 1924 1958	
FT	REPEAT 2004 2038	
FT	VARIANT 44 44	
FT	VARIANT 340 340	
FT	VARIANT 377 377	
FT	VARIANT 407 407	
FT	VARIANT 661 661	
FT	VARIANT 739 739	
FT	VARIANT 1038 1038	
FT	VARIANT 1198 1198	
FT	VARIANT 1257 1257	
FT	VARIANT 1392 1392	
FT	VARIANT 1520 1521	
FT	VARIANT 1583 1583	
FT	VARIANT 1613 1613	
FT	VARIANT 1686 1686	
FT	VARIANT 1799 1799	
FT	VARIANT 1881 1881	
FT	VARIANT 1894 1894	
FT	VARIANT 2141 2141	
FT	VARIANT 2392 2392	
FT	VARIANT 2605 2605	
FT	VARIANT 2648 2648	
FT	VARIANT 2717 2717	
FT	VARIANT 2729 2729	
FT	VARIANT 2814 2814	

-1- SUBUNIT: INTERACTS WITH RAD51 (BY SIMILARITY).  
-1- TISSUE SPECIFICITY: WIDELY EXPRESSED. HIGHEST EXPRESSION IN CEREBELLUM, TESTIS, ILEUM, APPENDIX, EPIDIDYMIS, OVARY AND MAMMARY GLAND. NO EXPRESSION IN LONG.  
-1- DEVELOPMENTAL STAGE: IN THE MAMMARY GLAND, EXPRESSION INCREASES DRAMATICALLY DURING PREGNANCY.  
-1- SIMILARITY: Contains 7 BRC2 repeats.  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

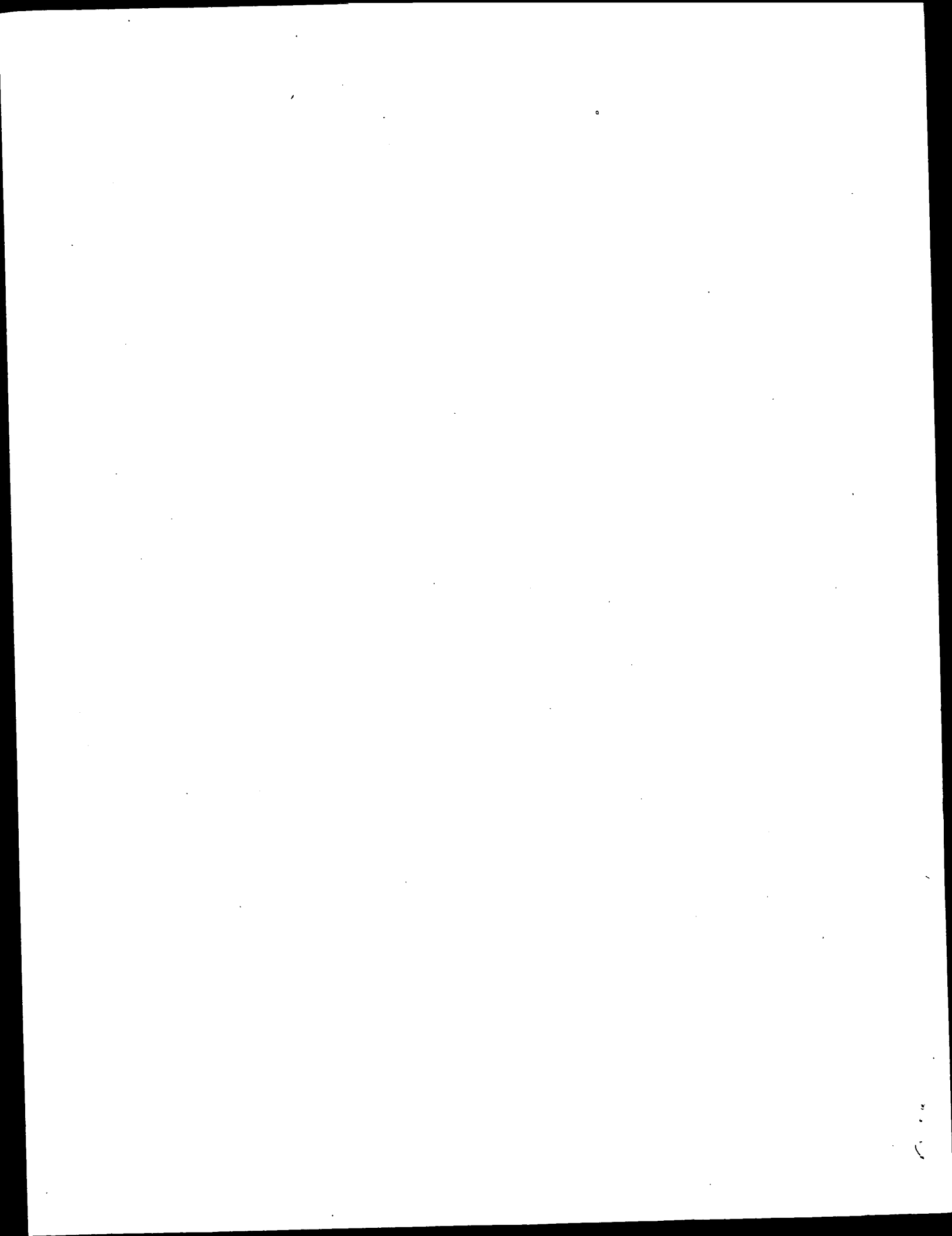
EMBL; U82270; AAB48306.1; -  
EMBL; U72947; AAB40720.1; -  
EMBL; U65594; AAC23702.1; -  
EMBL; U89652; AAB71377.1; -  
EMBL; U89503; AAC53276.1; -  
PIR; T30904; T30904  
PDB; 1MIU; 25-SEP-02.  
MGI:109337; Brc2.  
GO; GO:0005737; C:cytoplasm; IDA.  
GO; GO:0005634; C:nucleus; ISS.  
GO; GO:0005716; C:synaptonemal complex; ISS.  
GO; GO:0005515; F:protein binding activity; ISS.  
GO; GO:0003697; F:single-stranded DNA binding activity; ISS.  
GO; GO:0003653; F:transcriptional activator activity; ISS.  
GO; GO:0006915; P:apoptosis; ISS.  
GO; GO:0006338; P:chromatin modeling; ISS.  
GO; GO:0006260; P:DNA replication; ISS.  
GO; GO:0000724; P:double-strand break repair via homologous r. . .; ISS.  
GO; GO:0006325; P:establishment and/or maintenance of chromatin. . .; ISS.  
GO; GO:0007093; P:mitotic checkpoint; ISS.  
GO; GO:0007090; P:regulation of S phase of mitotic cell cycle; ISS.  
GO; GO:0045449; P:regulation of transcription; ISS.  
InterPro; IPR002093; BRC2\_repeat.  
Pfam; PF00634; BRC2; 7.  
PROSITE; PS0138; BRC2\_REPEAT; 6.  
POLYMERISM: Repeat; 3D-structure.  
REPEAT 981 1015  
REPEAT 1192 1226  
REPEAT 1394 1428  
REPEAT 1491 1525  
REPEAT 1623 1657  
REPEAT 1924 1958  
REPEAT 2004 2038  
VARIANT 44 44  
VARIANT 340 340  
VARIANT 377 377  
VARIANT 407 407  
VARIANT 661 661  
VARIANT 739 739  
VARIANT 1038 1038  
VARIANT 1198 1198  
VARIANT 1257 1257  
VARIANT 1392 1392  
VARIANT 1520 1521  
VARIANT 1583 1583  
VARIANT 1613 1613  
VARIANT 1686 1686  
VARIANT 1799 1799  
VARIANT 1881 1881  
VARIANT 1894 1894  
VARIANT 2141 2141  
VARIANT 2392 2392  
VARIANT 2605 2605  
VARIANT 2648 2648  
VARIANT 2717 2717  
VARIANT 2729 2729  
VARIANT 2814 2814

S -> F (IN STRAINS C57BL/6 AND 129/SV).  
T -> P (IN STRAIN 129/SV).  
N -> H (IN STRAIN C57BL/6).  
H -> P (IN STRAIN C57BL/6).  
I -> V (IN STRAIN C57BL/6).  
P -> H (IN STRAIN C57BL/6).  
I -> L (IN STRAINS C57BL/6 AND 129/SV).  
GF -> RI (IN STRAIN C57BL/6).  
Q -> P (IN STRAIN C57BL/6).  
Q -> R (IN STRAIN C57BL/6).  
FD -> CG (IN STRAIN C57BL/6).  
R -> W (IN STRAIN C57BL/6).  
C -> W (IN STRAIN C57BL/6).  
C -> W (IN STRAIN C57BL/6).  
S -> F (IN STRAIN C57BL/6).  
S -> F (IN STRAIN 129/SV).  
P -> L (IN STRAIN C57BL/6).  
Q -> F (IN STRAIN 129/SV).  
Q -> K (IN STRAIN C57BL/6).  
S -> R (IN STRAIN C57BL/6).  
K -> Q (IN STRAIN C57BL/6).  
A -> P (IN STRAIN C57BL/6).  
R -> C (IN STRAIN 129/SV).  
L -> M (IN STRAIN 129/SV).  
Q -> H (IN STRAIN C57BL/6).



Search completed: August 22, 2003, 18:56:42  
Job time : 28 secs

[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: August 22, 2003, 18:53:10 ; Search time 101 Seconds  
(without alignments)  
597.864 Million cell updates/sec

Title: US-10-030-613-1  
Perfect score: 1254  
Sequence: 1 MRSPGQGHVPEDGGLFLC.....SPKPKIOTRRGLPPSVSN 234

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues 830525

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23.\*  
1: sp-archaea.\*  
2: sp-bacteria.\*  
3: sp-fungi.\*  
4: sp-human.\*  
5: sp-invertebrate.\*  
6: sp-mammal.\*  
7: sp-mhc.\*  
8: sp-organelle.\*  
9: sp-phage.\*  
10: sp-plant.\*  
11: sp-rodent.\*  
12: sp-virus.\*  
13: sp-vertebrate.\*  
14: sp-unclassified.\*  
15: sp-rvirus.\*  
16: sp-bacteriaph.\*  
17: sp-archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	%	Query Match	Length	DB	ID	Description
1	1241	99.0	243	4	Q8N116	Q8N116	homo sapien
2	1188	94.7	243	11	Q8CH26	Q8CH26	rattus norv
3	1184	94.4	243	11	Q8VIP4	Q8VIP4	mus musculus
4	903	72.0	197	11	Q8CH27	Q8CH27	rattus norv
5	891	71.1	192	11	Q8VIP5	Q8VIP5	mus musculus
6	634	50.6	199	11	Q8K4S2	Q8K4S2	rattus norv
7	622.5	49.6	198	11	Q91WQ4	Q91WQ4	mus musculus
8	146	11.6	315	3	Q9P5S9	Q9P5S9	neurospora
9	126.5	10.1	249	3	Q9P4A1	Q9P4A1	cryptococcus
10	98	7.8	504	4	Q8IUY9	Q8IUY9	homo sapien
11	96	7.7	277	3	Q13703	Q13703	schizosacch
12	96	7.7	455	11	Q920L8	Q920L8	mus musculus
13	96	7.7	535	11	Q9EQT7	Q9EQT7	mus musculus
14	94	7.5	746	4	Q9BU60	Q9BU60	homo sapien
15	94	7.5	813	16	Q9X2G3	Q9X2G3	thermotoga
16	94	7.5	1021	4	Q15451	Q15451	homo sapien

17	94	7.5	1052	4	Q96FT1	Q96FT1	homo sapien
18	94	7.5	1130	4	Q8I2L8	Q8I2L8	homo sapien
19	94	7.5	1284	4	Q15450	Q15450	homo sapien
20	93.5	7.5	608	11	Q8BIG8	Q8BIG8	mus musculus
21	93.5	7.5	732	11	Q8K2L2	Q8K2L2	mus musculus
22	93.5	7.5	751	16	Q8YUK6	Q8YUK6	anabaena sp
23	92	7.3	270	5	Q93194	Q93194	caenorhabdi
24	92	7.3	649	10	Q9M6E6	Q9M6E6	nicotiana t
25	90.5	7.2	220	10	Q9FGA3	Q9FGA3	arabidopsis
26	90.5	7.2	339	10	Q8LJ34	Q8LJ34	oryza sativ
27	90.5	7.2	351	10	Q9FGF8	Q9FGF8	arabidopsis
28	90	7.2	479	10	Q9M6E4	Q9M6E4	nicotiana t
29	90	7.2	1066	4	Q9H2N5	Q9H2N5	homo sapien
30	90	7.2	3310	4	Q9P222	Q9P222	homo sapien
31	90	7.2	4025	4	Q9NR13	Q9NR13	homo sapien
32	90	7.2	4911	4	Q8NEZ4	Q8NEZ4	homo sapien
33	88.5	7.1	501	5	Q9BKQ7	Q9BKQ7	caenorhabdi
34	88	7.0	498	10	Q9MAD6	Q9MAD6	arabidopsis
35	88	7.0	513	5	Q95Q27	Q95Q27	caenorhabdi
36	88	7.0	1015	10	Q9SRH8	Q9SRH8	arabidopsis
37	88	7.0	1090	10	Q8LJ33	Q8LJ33	oryza sativ
38	88	7.0	3329	11	Q8VHD0	Q8VHD0	mus musculus
39	87.5	7.0	1557	5	Q96652	Q96652	drosophila
40	87	6.9	671	10	Q9FXA2	Q9FXA2	arabidopsis
41	87	6.9	862	2	Q9EXA5	Q9EXA5	neisseria m
42	86.5	6.9	1112	2	Q8GD29	Q8GD29	bordetella
43	86	6.9	528	2	Q9RLF1	Q9RLF1	pleistomonas
44	85	6.8	243	12	Q67788	Q67788	human adeno
45	85	6.8	408	11	Q8BYE5	Q8BYE5	mus musculus

ALIGNMENTS

RESULT 1

Q8N116 PRELIMINARY; PRT; 243 AA.  
ID Q8N116; AC Q8N116;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DE Calciineurin inhibitor ZAKI-4, beta splice variant 1 (Calcineurin  
inhibitor ZAKI-4 beta splice variant 2).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX PubMed=12102656;  
RA Cao X., Kambe F., Miyazaki T., Sarkar D., Ohmori S., Seo H.;  
RT "Novel human ZAKI-4 isoforms: Hormonal and tissue-specific regulation  
and function as calcineurin inhibitors.";  
RL Biochem. J. 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Cao X., Kambe F., Miyazaki T., Ohmori S., Seo H.;  
RT "Novel ZAKI-4 transcripts generated by alternative initiation and  
splicing from a single gene on chromosome 6.";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY034085; AAK59805.1;  
DR EMBL; AY034086; AAK59806.1;  
SQ SEQUENCE 243 AA; 27247 MW; 780073AE29A973E4 CRC64;

Query Match 99.0%; Score 1241; DB 4; Length 243;  
Best Local Similarity 99.1%; Pred. No. 5.8e-114;  
Matches 232; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MRSPGQGHVPEDGGLFLC...IDRDWAVTRCFEAFAQAITDFNDLPNSLFCACNVHQS...VF 60  
Db 10 MRSPGQGHVPEDGGLFLC...IDRDWAVTRCFEAFAQAITDFNDLPNSLFCACNVHQS...VF 69

QY 61 EGEESKEFEGFLRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 120  
 DB 70 EGEESKEFEGFLRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 129  
 QY 121 YFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVGVQPINATPVLNYDLYAVAKLG 180  
 DB 130 YFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVGVQPINATPVLNYDLYAVAKLG 189  
 QY 181 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 234  
 DB 190 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 243

## RESULT 2

Q8CH26 ID Q8CH26 PRELIMINARY; PRT; 243 AA.  
 AC Q8CH26; 2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE ZAKI-4 beta.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Brain;  
 RA Miyazaki T., Cao X., Kambe F., Ohmori S., Seo H.;  
 RT "Cloning of rat ZAKI-4 beta cDNA."  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF459023; AA015541.1; -  
 SQ SEQUENCE 243 AA; 27311 MW; 00E05BD1DCD0492D CRC64;

Query Match 94.7%; Score 1188; DB 11; Length 243;  
 Best Local Similarity 94.4%; Pred. No. 9.5e-109;  
 Matches 221; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 MRSFGQGHVPEDGGFLFLLCCIDRDWAVTRCFEAFAQATTFDNLPSLFACNVHQSVE 60  
 DB 10 MRSFGQGHVPEDGGFLFLLCCIDRDWAVTRCFEAFAQATTFDNLPSLFACNVHQSVE 69  
 QY 61 EGEESKEFEGFLRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 120  
 DB 70 EGEESKEFEGFLRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 129  
 QY 121 YFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVGVQPINATPVLNYDLYAVAKLG 180  
 DB 130 YFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVGVQPINATPVLNYDLYAVAKLG 189  
 QY 181 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 234  
 DB 190 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 243

## RESULT 3

Q8VIP4 ID Q8VIP4 PRELIMINARY; PRT; 243 AA.  
 AC Q8VIP4; 2003 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 23, Last annotation update)  
 DE Calcineurin inhibitory protein ZAKI-4 beta.  
 GN DSCRLL1 OR ZAKI-4 BETA.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kanou Y., Miyazaki T., Seo H., Murata Y.;  
 RT "Calcineurin inhibitory protein ZAKI-4."

Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RL EMBL: AB061525; BAB71956.1; -  
 DR MGD; MGI:1858219; Dscr111  
 DR InterPro; IPR006931; Calciopressin.  
 DR Pfam; PF04847; Calciopressin; 1.  
 DR Pfam; PF00076; rim; 1.  
 SQ SEQUENCE 243 AA; 27332 MW; 2CD8C1A810291851 CRC64;

Query Match 94.4%; Score 1184; DB 11; Length 243;

Best Local Similarity 94.0%; Pred. No. 2.3e-108;  
 Matches 220; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 1 MRSFGQGHVPEDGGFLFLLCCIDRDWAVTRCFEAFAQATTFDNLPSLFACNVHQSVE 60  
 DB 10 MRSFGQGHVPEDGGFLFLLCCIDRDWAVTRCFEAFAQATTFDNLPSLFACNVHQSVE 69  
 QY 61 EGEESKEFEGFLRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 120  
 DB 70 EGEESKEFEGFLRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIELHETQFRGKKLKL 129  
 QY 121 YFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVGVQPINATPVLNYDLYAVAKLG 180  
 DB 130 YFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVGVQPINATPVLNYDLYAVAKLG 189  
 QY 181 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 234  
 DB 190 PGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPGLPSPSVN 243

## RESULT 4

Q8CH27 ID Q8CH27 PRELIMINARY; PRT; 197 AA.  
 AC Q8CH27; 2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE ZAKI-4 alpha.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Brain;  
 RA Miyazaki T., Cao X., Kambe F., Ohmori S., Seo H.;  
 RT "Cloning of rat ZAKI-4 alpha cDNA."  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF459022; AA015540.1; -  
 SQ SEQUENCE 197 AA; 21991 MW; D8C906B7A7F46F4B CRC64;

Query Match 72.0%; Score 903; DB 11; Length 197;

Best Local Similarity 90.9%; Pred. No. 8e-81; Indels 0; Gaps 0;  
 Matches 170; Conservative 9; Mismatches 8;

QY 48 NSLFACNVHQSVEGEESKEFEGFLRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIE 107  
 DB 11 STLACVVDVEFTNQEVKEKFEGLFRTYDECVTQLFKSFRRVRINFSNPKSAARARIE 70  
 QY 108 LHETQFRGKKLKYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVGVQPINATPV 167  
 DB 71 LHETQFRGKKLKYFAQVQTPETDGDGKLHAPPQAKQFLISPPSSPPVGVQPINATPV 130  
 QY 168 LNYDLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPG 227  
 DB 131 LNYDLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDIEEEDPKTSPPKIIQTRRPG 190  
 QY 228 LPPSVSN 234  
 DB 191 LPPSVSN 197

## RESULT 5

[illegible]

```

RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL355932; CAB91442.1; -
DR InterPro; IPR006931; Calcipressin.
DR Pfam; PF04847; Calcipressin; 1.
SQ SEQUENCE 315 AA; 34194 MW; 767929876593C3AB7 CRC64;

Query Match 11.6%; Score 146; DB 3; Length 315;
Best Local Similarity 31.0%; Pred. No. 4.6e-06;
Matches 49; Conservative 21; Mismatches 60; Indels 28; Gaps 7;

QY 42 DFNDLP-----NSIFACNVHQ-SVFEGEESKEFGLFRYDDCV--TFOLF--- 85
DB 48 DLSNLPPLSQPTPPSNTLIFTNINSLDVSADN-----LQTRIDLSQTAFPAWSP 99
QY 86 -KSFRRVRINFENPKSAARARLHETQFRGKLLKLYFAQVQTPETDGDGKLHLAPPQPAK 144
DB 100 LKSFRRIRVTFDEQAATIAVRVDGEAILGERCHRVFGQTPPIDVSAADKHLALPDAGK 159
QY 145 QELISPPSPPPVGV--QPINDA--TPVLNYDLLYAVAKL 179
DB 160 LFFISPPSPPPHDWEQRMEDAPNTWVHAEDLAEALAKL 197

RESULT 9
Q9P4A1 PRELIMINARY; PRT; 249 AA.
AC Q9P4A1
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
GN CBP1.
OS Cryptococcus neoformans var. neoformans.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=40410;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H99;
RX MEDLINE=20359261; PubMed=10899116;
RA Gollach J., Fox D.S., Cutler N.S., Cox G.W., Perfect J.R., Heitman J.;
RT "Identification and characterization of a highly conserved calcineurin-
binding protein, CBP1/calcipressin, in Cryptococcus neoformans.";
RL EMBL J. 19; 3618-3629(2000).
DR EMBL; AF230799; AAF63734.1; -
DR InterPro; IPR006931; Calcipressin.
DR Pfam; PF04847; Calcipressin; 1.
SQ SEQUENCE 249 AA; 27293 MW; 78757183FA5EBF09 CRC64;

Query Match 10.1%; Score 126.5; DB 3; Length 249;
Best Local Similarity 38.0%; Pred. No. 0.00028;
Matches 30; Conservative 12; Mismatches 28; Indels 9; Gaps 2;

QY 118 LKLYFAQVQTPETDGDGKLHLAPPQPAKQFLISPPSPVGVQWQINDATP-----VLNYDLLY 174
DB 107 LRLHLVPTTLNPDPAATHLAPPPLPHNLISPPGSPGEGWPAEAPNRILLPEDLQR 166
QY 175 AVAKLPGGEKYLHAGTES 193
DB 167 AL-----ETLEINSGSKA 179

RESULT 10
Q8IU9 PRELIMINARY; PRT; 504 AA.
ID Q8IU9
AC Q8IU9;
DT 01-MAR-2003 (Tremblrel. 23, Created)

```

```

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to yes-associated protein, 65 kDa.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Pancreas;
RC Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038235; AAH38235.1; - 6145F7049ED338AE CRC64;
SQ SEQUENCE 504 AA; 54461 MW; 6145F7049ED338AE CRC64;

Query Match 7.8%; Score 98; DB 4; Length 504;
Best Local Similarity 31.1%; Pred. No. 0.45;
Matches 33; Conservative 11; Mismatches 30; Indels 32; Gaps 5;

QY 139 PPQPAKQFLISPPSPVGVQWQINDATPVLNYDLLYAVAKLPGGEKYLHAGTESPPSV- 197
DB 8 PPQPAQOGQGPSPQPPQGGPPS-----GPGQ--PAPAATQAAPQAP 48
QY 198 -----VVHV--CDSDIEEE-----DPKTSPPKPIIQTTRRGLPPS 231
DB 49 PAGQIVHVRGDSQTDLEALFNAMNPKTANVPQTPVPMRLKLPDS 94

RESULT 11
Q13703 PRELIMINARY; PRT; 277 AA.
ID Q13703
AC Q13703;
DT 01-JAN-1999 (Tremblrel. 09, Created)
DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein.
GN SPAC13F5.04C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Brown D., Churcher C.M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z99091; CAB11767.1; -
DR GeneDB SPombe; SPAC13F5.04c; -
DR Hypothetical protein.
RW SEQUENCE 277 AA; 540CF08359A06546 CRC64;
SQ SEQUENCE 277 AA; 540CF08359A06546 CRC64;

Query Match 7.7%; Score 96; DB 3; Length 277;
Best Local Similarity 22.4%; Pred. No. 0.32;
Matches 46; Conservative 22; Mismatches 77; Indels 60; Gaps 7;

QY 79 CVTFOLFSPKRRVRINF-----SNPKSAAR-----ARIELHETQFRGK 116
DB 2 CFFERLVSSFAILRISFLSITKIKFCKLRALQVAPQKAKTPSNHATELQOQSSTNST 61
QY 117 KLKLYFAQVQTPETDGDGKLHLAPP--QPAKQFLISPPSPVGVQWQ-----INDA 164
DB 62 TLPTQEAIVETNASASHETSFALPTTSPAASLSISPTKSAVSEPNVADVKSLSSTPA 121
QY 165 TPVLNYDLLYAVAKLPGGEKYE-----LHAGTESPSV-----VVHVCDSDIEE 208
DB 122 APQLN-----SPSHSYEPTTFPTTSTITENLPTIDTRTRSSSHIQSUSPESKQ 171

```

```

DR MGD; MGI:1927803; Fcamr.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000508; SigTase.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; Ig; 1.
DR PROSITE; PS00501; SPASE_I_1; 1.
KW Receptor.
SQ SEQUENCE 535 AA; 57696 MW; B275B77C70151C75 CRC64;

Query Match 7.7%; Score 96; DB 11; Length 535;
Best Local Similarity 25.0%; Pred. No. 0.76;
Matches 29; Conservative 19; Mismatches 50; Indels 18; Gaps 4;

QY 124 QVQTPETDGDK-----LHLAPPQAKQFLISPPS---SPPVGWQPINDATPVLNVDLLYAV 176
Db ||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
323 QGTPETDGRDETDVVRVPEAPRKTTGTRPSALISEHVYTWETLQDKTEVSKQQLHSL 382
QY 177 AKLGPGKEYE-----LHAGTSTPSVVVHVCDSDIEEDPKTSPKPKIIQTRRP 226
Db :||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
393 EELSPAPSAQPLNATCLEASEGRSI-----DCSLENTTEESSPTPSQLSVAGP 433

RESULT 14
Q9BU60 PRELIMINARY; PRT; 746 AA.
AC Q9BU60;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002875; AAH02875.1; -.
DR InterPro; IPR002965; P_rich_extensn.
DR PRINTS; PR01217; PRICHEXTENS.
DR Hypothetical protein.
KW NON_TER
FT 1
SQ SEQUENCE 746 AA; 79187 MW; 35A186C5A7ADEA8B CRC64;

Query Match 7.5%; Score 94; DB 4; Length 746;
Best Local Similarity 23.7%; Pred. No. 1.9;
Matches 41; Conservative 20; Mismatches 80; Indels 32; Gaps 6;

QY 61 EGEESEKEPGLFRYVDCVTFQLFKSFRRVIRNFSNPKSAARARIELHETQFRGKKLKL 120
Db ||| :||| : : : : : : : : : : : : : : : : : : : :
539 EEEEEEEFEFEFEFG-----ELEEEEEDEEEEEELEEVE----- 577
QY 121 YFAQVQTPETDGDKHLHAPPQAKQFLISPPSPVGMQPINDATPVLNVDLLYAVAKLG 180
Db : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
578 ---DLFEFGTAGVEEGAPPPPTLPALPPESPP-KVQPPPEPEP-----GLLEVEE-- 627
QY 181 PGEKVELHAGTSTPSVVVHVCDSDIEE-EDPKTSPKPKIIQTRRGLPPSPV 232
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
628 PCTEERGADTAPTLAEPALPSQGEVEREGESPAAGPPQELVLEEPSAPPTL 680

RESULT 15
Q9X2G3 PRELIMINARY; PRT; 813 AA.
AC Q9X2G3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cellobiose-phosphorylase.
GN TM1848.

```

Tue Aug 26 09:11:04 2003

```
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RL genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL; AE001822; AAC36910.1; -.
DR TIGR; TM1848; -.
DR InterPro; IPR000169; SHprot_acsite.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KW Complete proteome.
SQ SEQUENCE 813 AA; 93499 MW; 55FC07E7329D8C41 CRC64;

Query Match 7.5%; Score 94; DB 16; Length 813;
Best Local Similarity 24.5%; Pred. NO. 2.1; Indels 34; Gaps 9;
Matches 45; Conservative 27; Mismatches 78;

QY 51 FACNVHQSVPGEESKEKFEGLFRTYDDCVTFOL--FKSFRVRVINFENPKSAARARIEL 108
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
126 FTGEVHYLVLENKAERPKIKLFSEFIECLWNALDDMTNFQR---NYSTGEVEIEGVIY 182
QY 109 HETQFRGKKLKLIFPAQVQTP-----ETDGDK-LHLAPPQPAKOFLL--SPSSPPVGMQPI 161
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
183 HKTEYRERNHYAFYSVNQPIDGFDYDRESFGLYSGFEAPQAVVEGKPNVSGWAPI 242
QY 162 NDATPVLNYDLLYAVAKLGPGEKYEHLHAGTSTPSVVVHVHVCSDIEEEEDPKTSPKPII 221
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
243 ASH-----YLEIELAPSEKKEL-----IFIL-GYVENPEEEKWE-KPGVI 280
QY 222 QTRR 225
DB | |
281 NKRR 284
```

Search completed: August 22, 2003, 18:58:32  
Job time : 106 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2003, 07:02:29 ; Search time 376 Seconds  
(without alignments)  
6705.519 Million cell updates/sec

Title: US-10-030-613-3  
Perfect score: 934  
Sequence: 1 agaatcatacttcacggaa.....gtataaccattatttagaa 934

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	934	100.0	934	22	AAF25338
2	934	100.0	3159	24	ABA91457
3	720.8	77.2	3184	24	AAD30157
4	720.8	77.2	3184	24	ABL61768
5	720.8	77.2	3184	24	ABL65187
6	422	45.2	1021	24	ABA91458
7	413.2	44.2	594	24	AAD30154
8	279.2	29.9	828	24	AAD30158

9	257.6	27.6	615	24	ABA91463
10	231.6	24.8	720	24	AAD30159
11	211	22.6	597	24	AAD30152
12	209.6	22.4	597	24	AAD30153
13	197.6	21.2	599	24	AAD30151
14	194.4	20.8	2212	24	AAD30156
15	187	20.0	2331	24	AAD30155
16	187	20.0	2346	24	ABT11031
17	187	20.0	2355	25	ABX63255
18	184.2	19.7	2348	24	AAS94805
19	181.6	19.4	2358	21	AAF18328
20	177.4	19.0	2173	24	ABN96843
21	177.4	19.0	2174	20	AAO01282
22	167	17.9	363	22	AAI88711
23	133	14.2	227	21	AAI11239
24	126.8	13.6	531	24	ABA91461
25	116	12.4	1820	23	ABL29301
26	116	12.4	12550	23	ABL29300
27	93.2	10.0	442	21	AAC01774
28	83.6	9.0	365	24	ABS69673
29	83.6	9.0	412	22	ABA56172
30	83.6	9.0	412	22	AAK04364
31	83.6	9.0	412	22	AAI14449
32	83.6	9.0	412	22	AAI35821
33	83.6	9.0	412	22	AAI04272
34	83.6	9.0	412	23	ABS29505
35	83.6	9.0	412	24	ABS04419
36	83.6	9.0	446	22	ABA3406
37	83.6	9.0	446	22	ABA53852
38	83.6	9.0	446	22	ABA23598
39	83.6	9.0	446	22	AAK02113
40	83.6	9.0	446	22	AAK27563
41	83.6	9.0	446	22	AAI12148
42	83.6	9.0	446	22	AAI33493
43	83.6	9.0	446	22	AAI02061
44	83.6	9.0	446	23	ABS27133
45	83.6	9.0	446	24	ABS02041

#### ALIGNMENTS

RESULT 1  
AAF25338  
ID AAF25338 standard; cDNA; 934 BP.  
XX  
AC AAF25338;  
XX

DT 30-APR-2001 (first entry)

XX Nucleotide sequence of a human detoxification protein.

Human; detoxification protein; DETX; cancer; leukemia; melanoma;  
adenocarcinoma; autoimmune disorder; inflammatory disorder;  
rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;  
psoriasis; ulcerative colitis; infection; cell proliferative disorder;  
actinic keratosis; arteriosclerosis; cirrhosis; hepatitis; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
XX CDS 20..724

XX sig\_peptide /tag= a  
XX product= "detoxification protein"  
XX 20..100  
XX /tag= b

XX WO200104305-A2.

XX 18-JAN-2001.

XX 06-JUL-2000; 2000MO-US18509.

XX

*Applied*

07-JUL-1999; 99US-0142678.  
(INCY-) INCYTE GENOMICS INC.  
Tang YT, Yue H;  
WPI: 2001-147193/15.  
P-PSDB; AAB31788.  
New human detoxification protein and polynucleotide, useful for diagnosis, prevention and treatment of autoimmune/inflammatory disorders and cell proliferative disorders including cancer -  
Claim 5; Page 79; 79pp; English.  
The present sequence encodes a human detoxification polypeptide (DETX). DETX and its (ant)agonists are useful for preventing or treating disorders associated with decreased or increased expression or activity of DETX. DETX polypeptides are useful for screening compounds that specifically binds to DETX and for identifying (ant)agonists. Diseases prevented, treated and diagnosed include cancers (e.g. leukaemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow, breast, kidney, liver, pancreas, prostate and uterus), autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma, atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative colitis), bacterial, fungal, parasitic infections and cell proliferative disorders (e.g. actinic keratosis, arteriosclerosis, cirrhosis and hepatitis). Anti-DETX antibodies may be used as antagonists, as a targeting or delivery mechanism for bringing pharmaceutical agents into contact with cells or tissues expressing DETX and for diagnosis of DETX-related disorders.  
Sequence 934 BP: 247 A; 244 C; 206 G; 237 T; 0 other;  
Query Match 100.0%; Score 934; DB 22; Length 934;  
Best Local Similarity 100.0%; Pred. No. 4.4e-294;  
Matches 934; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAATCATACTTCGGAATGAGGAGCCAGCGACAGGACAGCGTCCCTGAAGATGG 60  
DB 1 AGAATCATACTTCATCGGAATGAGGAGCCAGCGACAGGACAGCGTCCCTGAAGATGG 60  
QY 61 AGGACTTTTCTTACTGTGCTGCATAGACAGGAGCTGGGCTGTCACCTGTTGTTGCAGA 120  
DB 61 AGGACTTTTCTTACTGTGCTGCATAGACAGGAGCTGGGCTGTCACCTGTTGTTGCAGA 120  
QY 121 AGAAGCCTTTCACGAATCACTGACTTCAATGACCTCCCAACTCGTGTGTTGCGTGCAA 180  
DB 121 AGAAGCCTTTCACGAATCACTGACTTCAATGACCTCCCAACTCGTGTGTTGCGTGCAA 180  
QY 181 TGTTCAACCACTGTCAGTGTGTTGAAGGAGAGAGAGCAAGAAATTTGAGGGACTGTTTCG 240  
DB 181 TGTTCAACCACTGTCAGTGTGTTGAAGGAGAGAGAGCAAGAAATTTGAGGGACTGTTTCG 240  
QY 241 GACTTATGATGACTGTGAGTGTCCAGCTATTTAAGAGTTTCAGAGAGTGTCCTGATAAA 300  
DB 241 GACTTATGATGACTGTGAGGTTCCAGCTATTTAAGAGTTTCAGAGAGTGTCCTGATAAA 300  
QY 301 CTTTCAGCAATCCTAAATCTGCAGCCGAGCTAGGATAGAGCTTCATGAACCAATTCAG 360  
DB 301 CTTTCAGCAATCCTAAATCTGCAGCCGAGCTAGGATAGAGCTTCATGAACCAATTCAG 360  
QY 361 AGGGAAAAAATTAAGCTCTACTTTGACAGGTTCCAGACTCCAGAGACAGATGGAGACAA 420  
DB 361 AGGGAAAAAATTAAGCTCTACTTTGACAGGTTCCAGACTCCAGAGACAGATGGAGACAA 420  
QY 421 ACTGCACTTGGGTCCACCCCGCCTGCCAAACAGTTTCTCATCTCGCCCTTCCTCCC 480  
DB 421 ACTGCACTTGGGTCCACCCCGCCTGCCAAACAGTTTCTCATCTCGCCCTTCCTCCC 480  
QY 481 ACCTGTTGGCTGCGACCCCATCAACGATGCGCCAGTCCCTCAACTATGACCTCTCTA 540  
DB 481 ACCTGTTGGCTGCGACCCCATCAACGATGCGCCAGTCCCTCAACTATGACCTCTCTA 540

\_\_\_\_\_

AC ABL61768;

Claim 1; SEQ ID 105; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (1) of a signature gene set, where (1) comprises a sequence (S) selected from 8447 sequences (given in ABL611664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (1) has cytostatic activity and can be used in anti-neoplastic therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

Query Match	77.2%;	Score 720.8;	DB 24;	Length 3184;
Best Local Similarity	95.9%;	Pred. No. 4e-284;		
Matches 740;	Conservative 0;	Mismatches 32;	Indels 0;	Gaps 0;
QY	163	CTCGTTGTTGGCTGTCGAATTTTCACACGATCACTGTTTGAAGGAGAGACACAGAGAAA	222	
Db	222	CAC TCTGGTGGCTGTGTGGTGGATGTGCGAGCTCTTACC AATCAGAGGGTTTAAAGGAAA	281	
QY	223	ATTTGAGGACACTGTTTTCGGACTTATGATGACTGTGTGACGTTCCAGCTATTAAAGAGTTT	282	
Db	282	ATT TGGG GAC TGT TTCGGACTTATGATGACTGTGTGACGTTCCAGCTATTAAAGAGTTT	341	
QY	283	CAGACGTGCCGTATAAACTTCAGCAATCCTAAATCTGCAGCCGAGCTAGGATAGAGCT	342	
Db	342	CAGAGCTGTCCGTATAAACTTCAGCAATCCTAAATCTGCAGCCGAGCTAGGATAGAGCT	401	
QY	343	TCATCAAAACCAATTCAGAGGGAAAAATTAAAGCTCTACTTTGCACAGGTTTCAGACTCC	402	
Db	402	TCATGAACCCAAATTCAGAGGGAAAAATTAAAGCTCTACTTTGCACAGGTTTCAGACTCC	461	
QY	403	AGACAGATGGAGACAACTGCACTTGGCTTCACCCACGCTGCCAACAGTTTCTCAT	462	
Db	462	AGAGACAGATGGAGACAACTGCACTTGGCTTCACCCACGCTGCCAACAGTTTCTCAT	521	
QY	463	CTCGCCCTTCTCPCCCACCTGTTGGCTGGCAGCCCATCAAGATGCCACGCGCTCT	522	
Db	522	CTCGCCCTTCTCPCCCACCTGTTAGCTGGCAGCCCATCAAGATGCCACGCGCTCT	581	
QY	523	CAACTATGACCTCCTCTATGCTGTGCCAACTAGACACGAGAGAAGTATGAGCTCCA	582	
Db	582	CAACTATGACCTCCTCTATGCTGTGCCAACTAGACACGAGAGAAGTATGAGCTCCA	641	
QY	583	TGCAGGACATGATGCACCCCAAGTGTCTGTGTCAGCTGTGCGACATGACATAGAGA	642	
Db	642	TGCAGGACATGATGCACCCCAAGTGTCTGTGTCAGCTGTGCGACATGACATAGAGA	701	
QY	643	AGAAGAGGACCCAAAGACTTCCCAAAGCCAAAATATCCAACTCGCCGCTCTGGCCT	702	
Db	702	AGAAGAGGACCCAAAGACTTCCCAAAGCCAAAATATCCAACTCGCCGCTCTGGCCT	761	
QY	703	GCACCCCTCCGTGTCACATGAGCTGCCTTCTCTCGATATAGCCGCTCTCTTT	762	
Db	762	GCACCCCTCCGTGTCACATGAGCTGCCTTCTCTCGATATAGCCGCTCTCTTT	821	
QY	763	ATCATGCTTTTTCCTCCCTGTTGTTGTCAAAAAAATTCGCTTTTAAATTCCTGGGTGTTT	822	
Db	822	ATCATGCTTTTTCCTCCCTGTTGTTGTCAAAAAAATTCGCTTTTAAATTCCTGGGTGTTT	881	
QY	823	GTTTCTTTGAGATTCCTTCCCTTGTTATCAAGCCCTCTCGACAAAGGGCTAGGAAAAGGT	882	
Db	882	GTTTCTTTGAGATTCCTTCCCTTGTTATCAAGCCCTCTCGACAAAGGGCTAGGAAAAGGT	941	



```

Db 763 ATTTGAGGAGCTGTCGGACCTATGACGAATGTGTGACGCTTTCAGCTGTTAAAGAGTT 704
QY 283 CAGACGTG--TCCGTATAAATTCAGCAATCTTAAATCTGCAGCCG--AGCTAGATAGA 339
Db 703 CCGACGGTTCGGAATAAATTCAGCACCCCAAGCTGCAGCCCGTGCCTCGGATAGA 644
QY 340 GCTTCA--TGAACCAATTCAGAGGGGAAATTAAGAGCTCTACTTTGCACAGTTTCAG 397
Db 643 GCTTCATGTGGACCCAGTTCATGAGGGAAGAGCTGAAACTCTACTTTGCACAGTCCAG 584
QY 398 ACTCCAGACAGATGG--AGACAACTGCA--CTTGGCTTCCACCCAGCTGCCAAACAGT 455
Db 583 ACCCCAGACAGATGGTAGAACAATCGATTTGGCACCCCAAGCTGCCAAACAGT 524
QY 456 TTCTCATCTCGCCCTTCTCCCTCCACCTGTGGCTGGCAGCCCATCAACGATGCCAGC 515
Db 523 TCCTCATCTACCCCTTCTCTCTCTCCGTTGGCTGGAAGCTTATCAGGATGCCACAC 464
QY 516 CAGTCCCTAAC--TATGACCTCTCTATGCTGTGGCCAACTAGGACCAAGAGATAT 574
Db 463 CAGTCCCTAACGATAGGACCTCTCTATGCTGTGGCCAACTAGGACCAAGAGATAT 404
QY 575 GAGTCCATCAGGAGCTGAGTCCACCCCAAGTGTGCTGTCGACGTTGTCGACAGTGAC 634
Db 403 GAGTGCATCGGGAAGTCTACACCGAGGCTGTGTCGACGTTGTCGACAGTGAC 344
QY 635 ATAGAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 594
Db 343 TTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 284
QY 695 CTGGGCTGCCACCCCTCGTGTGCAACTGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 754
Db 283 CTGGGCTGCCACCCCTCGTGTGCAACTGAGTGA-----GCAGTCCATGATGATGATC 231
QY 755 TCCTCTTTATCATGCTTTTCCCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 813
Db 230 TCCTC--TTATCATGCTTTTCCCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 172
QY 814 TGGGTGTTGGTGTGTTGAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 872
Db 171 TGGGTGTTGGTGTGTTGAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 112
QY 873 AGGAAAGTGATATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 932
Db 111 AGGAAAGTGATATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 932
QY 933 AA 934
Db 51 AA 50

```

## RESULT 7

AAD30154

ID AAD30154

XX AAD30154 standard; DNA; 594 BP.

AC AAD30154;

XX AAD30154;

DT 17-MAY-2002 (first entry)

XX Mouse MCIP associated DNA #4.

DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

KW heart failure; cardiomyopathy; heart disease; mouse; gene; ds.

XX Mus musculus.

XX Key

FH Location/Qualifiers

FT 1..594

FT /\*tag= a

FT /product= "Mouse MCIP associated protein #4"

XX

XX

PN WO200204491-A2.

XX PD 17-JAN-2002.

XX PF 06-JUL-2001; 2001WO-US21662.

XX PR 07-JUL-2000; 2000US-216601P.

XX PR 13-FEB-2001; 2001US-0782953.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PA (WILL) WILLIAMS S R.

XX PA (ROTH) ROTHERMEL B.

XX PI Williams SR, Rothermel B;

XX XX WPI; 2002-179698/23.

XX DR P-PSDB; AAE18913.

XX XX

XX PT Screening for modulators of muscle calcineurin interacting protein

XX PT (MCIP) binding, expression or phosphorylation, useful for treating

XX PT cardiac hypertrophy or heart failure, comprises mixing MCIP,

XX PT calcineurin and a test compound.

XX PS Disclosure: Page 153-154; 174pp; English.

XX XX

XX CC The invention relates to muscle calcineurin interacting proteins (MCIPs)

XX CC and nucleic acid molecules encoding such proteins. MCIPs form a physical

XX CC complex with the catalytic subunit of calcineurin and increased levels

XX CC of MCIPs correspond to a reduced ability of calcineurin to stimulate

XX CC transcription of certain target genes. The invention also relates to

XX CC methods for identifying modulators of MCIP binding, expression or

XX CC phosphorylation. Inhibitors or promoters of MCIP binding, expression or

XX CC may be used for treating cardiac hypertrophy and heart failure.

XX CC Antibodies to MCIP can be used in characterizing the MCIP content of

XX CC healthy and diseased tissues and subsequently for determining the

XX CC presence or absence of cardiomyopathy or as predictor of heart disease.

XX CC The present sequence is mouse MCIP associated DNA.

XX CC Note: This sequence has been described as mouse MCIP1 in the

XX CC specification, however the sequence seems to be a polynucleotide

XX CC encoding a MCIP associated protein.

XX XX

XX SQ Sequence 594 BP; 148 A; 165 C; 149 G; 132 T; 0 other;

## Query Match

Best Local Similarity 44.2%; Score 413.2; DB 24; Length 594;

Matches 469; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 163 CTCGTTGTTGGCTGCAATGTTCCACGACGAGTGTGTTGAAGGAGAGAGAGAGAGAGAG 222

Db 33 CACTCTGGTCGCTGTGTGTGATGTGGAGGTCTTTACCAATCAGGAGGTAAAGGAAA 92

QY 223 ATTTGAGGAGCTGTTTCGGACCTATGATGACTGTGTGAGCTTTCAGCTATTTAAGAGTTT 282

Db 93 ATTCGAGGAGCTGTTCCGAGCCTATGATGAATGTGTGAGTTCAGCTGTTTAAAGAGTTT 152

QY 283 CAGAGCTGTCCGTATATAAATTCAGCAATCTTCAAGCAATCTTCAAGCAATCTTCAAGCAAT 342

Db 153 CCGACGGTTCGATATAAATTTTCAGCCATCTCCCAATCTCGACCCGTCGCGGATAGAGCT 212

QY 343 TCATGAAACCCCAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 402

Db 213 TCATGAGACTCAGTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 272

QY 403 AGAGACAGATGGAGACAAACTGACCTTGGCTGCCACCCAGCTGCCACCCAGCTGCCACCC 462

Db 273 AGAGACAGATGGAGACAAACTGACCTTGGCTGCCACCCAGCTGCCACCCAGCTGCCACCC 332

QY 463 CTGCCCCCTTCTCCACCTGTGTGGCTGGCAGCCCATCAAGATGCCAGCTGCCAGCTGCC 522

Db 333 CTACCCCTTCTATCTCTCTCTGTGGCTGAAGGCTATCAGCGATGCCACACAGCTCCT 392

QY 523 CAACATGACCT 582

Db 393 CAACATGACCT 452







us-10-030-613-3.rng

Tue Aug 26 09:11:04 2003

cardiac hypertrophy or heart failure, comprises mixing MCIP, calcineurin and a test compound -  
Claim 95; Page 147-148; 174pp; English.

The invention relates to muscle calcineurin interacting proteins (MCIPs) and nucleic acid molecules encoding such proteins. MCIPs form a physical complex with the catalytic subunit of calcineurin and increased levels of MCIPs correspond to a reduced ability of calcineurin to stimulate transcriptions of certain target genes. The invention also relates to methods for identifying modulators of MCIP binding, expression or phosphorylation. Inhibitors or promoters of MCIP binding, expression or phosphorylation may be used for treating cardiac hypertrophy and heart failure. Antibodies to MCIP can be used in characterizing the MCIP content of healthy and diseased tissues and subsequently for determining the presence or absence of cardiomyopathy or as predictor of heart disease. The present sequence is human MCIP associated DNA.

Sequence 597 BP; 160 A; 170 C; 151 G; 116 T; 0 other;

Query Match 22.6%; Score 211; DB 24; Length 597;

Best Local Similarity 65.2%; Pred. No. 5e-58; Indels 12; Gaps 1; Matches 332; Conservative 0; Mismatches 165;

143 GACTTCATGACCTCCCAACTCGTTGTTGGTGGCAATGTTACACAGTCACTGTTGAA 202  
13 GATCTGCGAGGACCTGCGGCGCCACCATCGCTGCCACCTGGACCGCGGTGTTG 72  
203 GGAGAGAGAGACAGAAATTTTCAGGAGCTGTTTCGGACTTATGATGACTGTGTGACG 262  
73 GACGCGCTGTGCGGGGCCAAATTTTGAATCCCTCTTCAGAACATATGACAAGGACACACC 132  
263 TTCCAGCTATTAAAGAGTTTCAGAGCTGTCCGTATATAAATCTTCAGCAATCTTAAATCTGCA 322  
133 TTCCAGTATTTAAGAGCTTCAACAGCTGTCCGGATAAAATTCAGCAACCCCTTATCTGCA 192  
323 GCCGAGCTAGGATAGAGCTTCATGAAACCCCAATTCAGAGGGAGAAATTAAGGCTCTAC 382  
193 GCCGATGCCAGGCTGCGGCTGCACAAGACCGAGTTCCTGGGAAGGAAATGAAGTTGTAT 252  
383 TTTTCACAGGTTTCAGAGCTCCAGAGACAGATGGAGACAAACTGCACCTGGCTCCACCCAG 442  
253 TTTGCTCAGACT-----TTACACATAGGAAGTTACACCTGGCTCGGCCCAAT 300  
443 CTTGCCAAACAGTTTCTCATCTCGCCCTTCCTCCCTCCCTGCTGGTGGAGCCCATC 502  
301 CCCGACAAACAGTTCTCTCATCTCCCTCCGCGCTCTCTCCCTGGCTGGAAACAGTA 360  
503 AACGATGCCAGCGGCTCCTCACTATGACCTCTCTCTATCTGCTGGCCAACTAGGACCA 562  
361 GAAGTGGCCACCCCGTCAATAAATAGCATCTTTTATATGCTATCTCCAGCTGGGGCCA 420  
563 GGAGAGAGATGAGCTCCATGCGAGGAGTCCAGTCCACCCCAAGTGTGCTGTCACGCTG 622  
421 GGAGAGAGATGAGCTCCATGCGAGGAGTCCAGTCCAGTCCACCCCACTCCAGTGTGTCACG 480  
623 TGCAGAGTGCATAGAGAGAGAGAGAGGA 651  
481 TGTGAGAGTGACCAAGAGAGATGAGGAGGA 509

RESULT 12  
AAD30153  
ID AAD30153 standard; DNA; 597 BP.

XX AAD30153;  
AC AAD30153;

DT 17-MAY-2002 (first entry).

XX Mouse MCIP associated DNA #3.

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
KW heart failure; cardiomyopathy; heart disease; mouse; gene; ds.

265 CCAGCTATTAAAGAGTTTCAGAGCTGTCCGTATATAAACTTCAGCAATCTTAAATCTGCAGC 324  
187 TCAGCTGTTTAAAGAGCTTTAGAGAGTTCAGAAATTAATTCAGCAACCTTGAAGCGCAGC 246  
325 CCGAGCTAGGATAGAGCTTCATGAAACCCCAATTCAGAGGAGAAATTAAGCTCTACTT 384  
247 AAGAGCGGATAGAGCTTCAGCAAGAACAGAGCTTCAATGGCGAGAGCTAAAGCTATATT 306  
385 TGCAGAGTTTCAGACTCCAGAGACAGATGGAGACAAACTGCACCTTGGCTCCACCCAGCC 444  
307 TGCAGAG-----TCTATCTCTCTGCGCGCCAGCC 336  
445 TCCCAAAACAGTTTCTCATCTCGCCCGCTTCCTCCCGCTGTTGGCTGGAGCCCATCAA 504  
337 TGTCAAGAGTTTCTCATCTCGCCCTCCAGCTTCCCGCTGAGTCCGAGTGGAGAGAGGA 396  
505 CGATGCCAGCGGCTCCTCAACTATGAGCTCTCTATGCTGTGGCCAACTAGACACAGG 564  
397 AGATGGGATGCTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 456  
565 AGAAGAGTATGAGCTCCATGAGGAGTCCAGTCCAGCCCAAGTGTGCTGCGTGCAGTGTG 624  
457 AGAAGAAATATGAATTCAGCGCGGAGACAGAGTCCGAGTCCAGCTGGTGGTTCATGCTG 516  
625 CGACAGTGCATAG 684  
517 TGAAGTGAAGTGAAG 573  
685 AACTCGGCGCTGCGCTGCCACCTCCGCTGTCCTCA 720  
574 GACAAGGCGCGCGGACCTCCGAGCGGCTTGA 609

RESULT 11  
AAD30152  
ID AAD30152 standard; DNA; 597 BP.  
XX AAD30152;  
AC AAD30152;  
DT 17-MAY-2002 (first entry)  
XX Human MCIP associated DNA #2.  
XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
KW heart failure; cardiomyopathy; heart disease; human; gene; ds.

XX Homo sapiens.  
XX key Location/Qualifiers  
XX CDS 1..597  
FT FT /tag= a  
FT FT /product= "Human MCIP associated protein #2"  
XX W0200204491-A2.  
XX 17-JAN-2002.  
XX 06-JUL-2001; 2001WO-US21662.  
XX 07-JUL-2000; 2000US-216601P.  
XX 13-FEB-2001; 2001US-0782953.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX (WILL/) WILLIAMS S R.  
XX (ROTH/) ROTHERMEL B.  
XX Williams SR, Rothermel B;  
XX WPI; 2002-179698/23.  
XX P-PSDB; AAE18911.  
XX Screening for modulators of muscle calcineurin interacting protein  
XX (MCIP) binding, expression or phosphorylation, useful for treating



Query Match 21.2%; Score 197.6; DB 24; Length 599;  
 Best Local Similarity 64.6%; Pred. No. 1.2e-53;  
 Matches 332; Conservative 0; Mismatches 169; Indels 13; Gaps 2;

QY 138 TCACTGACCTCAATGACCTCCCAACTCGTGTGCGTGCATGTTCCACGAGTCAGTGT 197  
 DB 87 TTAGGACCTTAGTCAATTTTAGTCCCTGCTGTTGCGTGCATGATGCT 146  
 QY 198 TTGAAGGAGAGACAGGAGAAATTTGAGGAGCTGTTCCGACTTATGATGACTGTG 257  
 DB 147 TCAGCGAAAGTGAGACAGGCGCAATTTGATCCCTCTTCAGACATATGACAGGACA 206  
 QY 258 TGACGTTCCAGCTATTAAAGATTTCAGAGCTGTCCGTATAAACTTCAGCAATCTTAAT 317  
 DB 207 CCACCTCCAGTATTAAAGAGCTTCAACAGCTGTCGGATAAACTTCAGCAACCCCTTAT 266  
 QY 318 CTGAGCCCGAGCTAGATAGAGCTTCATGAACCACTTCAGAGGAGGAAAAATTAAGC 377  
 DB 267 CTGAGCCGATGCCAGCTCGGCTGCACAGACCGAGTCTCTGGGAGGAATGAATGAAT 326  
 QY 378 TCTACTTTGACAGCTTCAGACTCCAGAGACAGATGGAGACAACTGCACCTGGCTCCAC 437  
 DB 327 TGTATTTGCTCAGACT-----TTACACATAGGAAGTTTCACACCTGGCTCC-G 373  
 QY 438 CCCAGCTGCCAAACAGTTCTCATCTCGCCCTTCTCCACCTGTTGGCTGGCAGC 497  
 DB 374 CCAATCCCGCAACACAGTCTCTCATCTCCCTCCGCTCTCTCCGCTGGCTGGNAAC 433  
 QY 498 CCATCAAGATGCGCAGCTGCTCAACTATGACCTCTCTATGCTGTGGCCAACTAG 557  
 DB 434 AAGTAGAGATGCCACCCCGCTCATAAATTAGCATCTTTATATGCCATCTCCAAGCTGG 493  
 QY 558 GACCAAGGAGAGATGATGAGTCCATGAGGAGTCTGAGTCCACCCCAAGTGTCTGCTGC 617  
 DB 494 GCGCAGGAGAGAGATGATGAATGCTGATGCGAGCAGCAGACACCACTCCCGAGTGTGCTGC 553  
 QY 618 AGCTGTGCGAGTGACATAGAGGAAGAGGA 651  
 DB 554 AGCTGTGAGAGTGACCAAGAGATGAGGAGGA 587

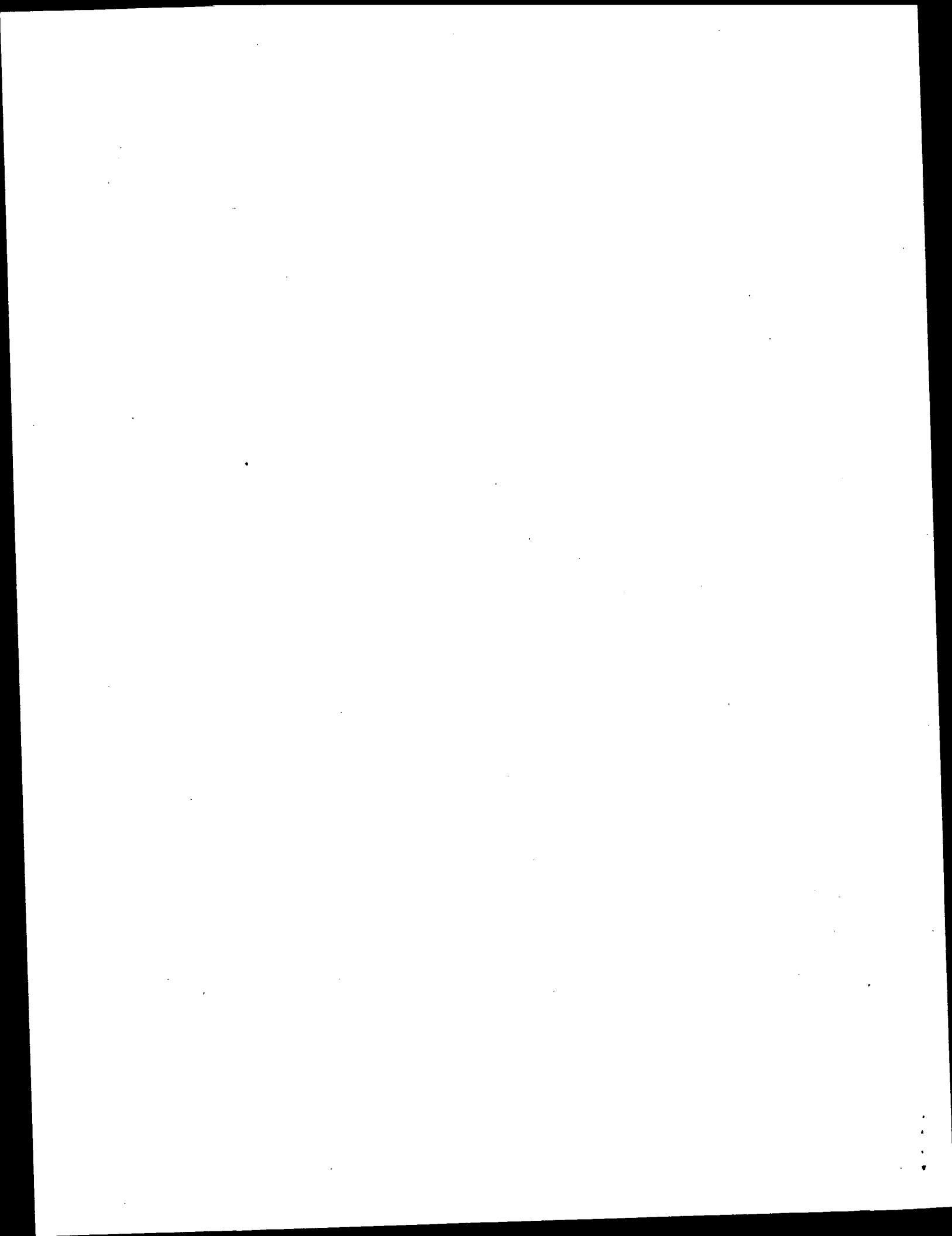
RESULT 14  
 AAD30156  
 ID AAD30156 standard; DNA; 2212 BP.  
 AC AAD30156;  
 XX 17-MAY-2002 (first entry)  
 DT Human MCIP associated DNA #2.  
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;  
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 25..618  
 FT /\*tag= a  
 FT /product= "Human MCIP associated protein #2"

WO200204491-A2.  
 17-JAN-2002.  
 06-JUL-2001; 2001WO-US21662.  
 07-JUL-2000; 2000US-216601P.  
 13-FEB-2001; 2001US-0782953.  
 (TEXA) UNIV TEXAS SYSTEM.  
 PA (WILL/) WILLIAMS S R.  
 PA (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;  
 XX WPI; 2002-179698/23.  
 DR P-PSDB; AAE18915.  
 XX Screening for modulators of muscle calcineurin interacting protein  
 PT (MCIP) binding, expression or phosphorylation, useful for treating  
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,  
 PT calcineurin and a test compound -  
 XX  
 PS Example 1; Page 159-161; 174pp; English.  
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)  
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical  
 CC complex with the catalytic subunit of calcineurin and increased levels  
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate  
 CC transcription of certain target genes. The invention also relates to  
 CC methods for identifying modulators of MCIP binding, expression or  
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin  
 CC may be used for treating cardiac hypertrophy and heart failure.  
 CC Antibodies to MCIP can be used in characterizing the MCIP content of  
 CC healthy and diseased tissues and subsequently for determining the  
 CC presence or absence of cardiomyopathy or as predictor of heart disease.  
 CC The present sequence is human MCIP associated DNA.  
 CC Note: This sequence has been described as murine MCIP splice variant in  
 CC the specification, however the sequence seems to be a polynucleotide  
 CC encoding a MCIP associated protein.

XX Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;  
 QY Query Match 20.8%; Score 194.4; DB 24; Length 2212;  
 DB Best Local Similarity 60.8%; Pred. No. 2.9e-52;  
 QY Matches 365; Conservative 0; Mismatches 211; Indels 24; Gaps 2;  
 DB 143 GACTTCAATGACCTCCCAACTCGTGTGTTGGTGTCAATGTTCCACGAGTCAGTGTGAA 202  
 QY 37 GACCTGCAGGACCTGCCAGCGCCACCATCGCTGTCACTGGACCGCGGTGTCGTG 96  
 DB 203 GGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 262  
 QY 97 GACGCGCTGTGCGCGGCGCAAAATTTAGTTCCTTTAGGACGTATGACAGGAGCATCAC 156  
 DB 263 TTCCAGCTATTAAAGAGTTTTCAGACGTGTCCGTATATAAATTCAGCAATCTTAATCTGCA 322  
 QY 157 TTTTCAGTATTAAAGAGCTTCAAAAGAGTCAAGATTAACATTCAGCAACCCCTTCTCCGCA 216  
 DB 323 GCCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAGAGGAGGAGGAGGAGGAGGAGGAG 382  
 QY 217 GCAGATGCCAGGCTCCAGCTGCATAGAGTGTCTGCGGAGGAGGAGGAGGAGGAGGAGGAG 276  
 DB 383 TTTCCAGAGTTTCAGACTCCAGAGACAGATGGAGACAAACTGCACTTTGGTTCACCCAG 442  
 QY 277 TTTGCTCAG-----ACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324  
 DB 443 CCTGCCAAACAGTTTCTCATCTCGCCCTTCTCCCTGCTGTCGAGCCCATC 502  
 QY 325 CCAGACAAAGCAGTTTCTGATCTCCCTCCCGCTCTCCGCGAGTGGATGGAACAGTG 384  
 DB 503 AACGATGCCAGCAGCTCTCAACTATGACCTCTATGCTGTGGGCAAACTAGGACCA 562  
 QY 385 GAAGATGCGACCCCACTCAATATGATCTCTTATATGCTATGCTATGCTATGCTATGCTATGCT 444  
 DB 563 GGAGAGAGATGAGCTCCATGAGGAGGAGTCCACCCCAAGTGTCTGTCACGCTG 622  
 QY 445 GGGGAAAAAGTATGAATTTGACGCGGAGTCCACACTCCAGCGGTGGTGTCCATGTA 504  
 DB 623 TCGGACACTGACATAGAG-----GAAGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 670  
 QY 505 TGTGAGAGTGTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 564  
 DB 671 CCAAAAAATCATCCAAACTCGGCGTCTTGGCTGCGCACCTCCGCTGTCCAACTGAGCTGCC 730  
 QY





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2003, 11:27:06 ; Search time 322 Seconds  
(without alignments)  
6521.071 Million cell updates/sec

Title: US-10-030-613-3

Perfect score: 934

Sequence: 1 agaatacattcttcgga.....gtataaccattatttagaa 934

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 151743 seqs, 1124081882 residues

Total number of hits satisfying chosen parameters: 3034486

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	720.8	77.2	3184	10	US-09-954-456-497
2	720.8	77.2	3184	10	US-09-782-953-17
3	413.2	44.2	594	10	US-09-782-953-8
4	279.2	29.9	828	10	US-09-782-953-20
5	231.6	24.8	720	10	US-09-782-953-23
6	211	22.6	597	10	US-09-782-953-2
7	209.6	22.4	597	10	US-09-782-953-5
8	197.6	21.2	599	10	US-09-782-953-1
9	194.4	20.8	2212	10	US-09-782-953-14
10	187	20.0	2331	10	US-09-782-953-11
11	187	20.0	2355	13	US-10-044-090-255
12	187	20.0	2358	14	US-10-084-817-36
13	181.6	19.4	2358	9	US-09-925-302-347
14	177.4	19.0	2173	10	US-09-880-107-3340
15	177.4	19.0	2173	14	US-10-102-524-1746
16	177.4	19.0	2409	14	US-10-198-846-13070

17	83.6	9.0	365	10	US-09-728-445-736
18	83.6	9.0	412	9	US-09-864-761-10388
19	83.6	9.0	446	9	US-09-864-761-2064
20	79.8	8.5	518	14	US-10-102-524-1126
21	64	6.9	3404	13	US-10-027-632-114005
22	56	6.0	111	9	US-09-864-761-27019
23	47.2	5.1	486	9	US-09-864-761-17592
24	46.4	5.0	123	9	US-09-864-761-18808
25	43.6	4.7	85	9	US-09-920-671-11
26	37.6	4.0	139257	11	US-09-764-891-6418
27	37.2	4.0	17904	11	US-10-205-428-708
28	37.2	3.8	671	14	US-10-184-644-346
29	35.4	3.8	671	14	US-10-184-644-346
30	35.4	3.8	873	13	US-10-027-632-154378
31	35.2	3.8	873	13	US-10-027-632-154378
32	35.2	3.8	589	13	US-10-027-632-225823
33	34.8	3.7	589	13	US-10-027-632-225823
34	34.8	3.7	590	13	US-10-027-632-211428
35	34.4	3.7	590	13	US-10-017-161-781
36	34.2	3.7	7972	12	US-10-027-632-191694
37	34	3.6	533	13	US-09-901-136-3
38	34	3.6	378361	11	US-10-027-632-115812
39	33.4	3.6	621	13	US-10-140-472-358
40	33.4	3.6	1049	12	US-10-141-761-358
41	33.4	3.6	1049	12	US-10-123-155-358
42	33.4	3.6	1049	14	US-10-146-731-358
43	33.4	3.6	1049	15	US-09-918-995-7591
44	33.2	3.6	361	11	US-09-918-995-7591
45	33.2	3.6	432	11	US-09-918-995-37189

## ALIGNMENTS

## RESULT 1

US-09-954-456-497  
; Sequence 497, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using  
; FILE REFERENCE: Sets  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; CURRENT FILING DATE: 2001-09-18  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 497  
; LENGTH: 3184  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-497

Sequence 736, App  
Sequence 10388, A  
Sequence 2064, Ap  
Sequence 1126, Ap  
Sequence 114005,  
Sequence 27019, A  
Sequence 809, App  
Sequence 17592, A  
Sequence 18808, A  
Sequence 11, Appl  
Sequence 6418, Ap  
Sequence 708, App  
Sequence 346, App  
Sequence 346, App  
Sequence 154378,  
Sequence 154379,  
Sequence 225823,  
Sequence 225824,  
Sequence 211428,  
Sequence 781, App  
Sequence 191694,  
Sequence 3, Appl  
Sequence 115812,  
Sequence 358, App  
Sequence 358, App  
Sequence 358, App  
Sequence 358, App  
Sequence 37189, A

```
; SEQ ID NO 17  
; LENGTH: 3184  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (205)..(780)  
uts-09-782-953-17
```

[illegible]

```

Query Match          77.2%; Score 720.8; DB 10; Length 3104;
Best Local Similarity 95.9%; pred. No. 1.4e-226;
Matches 740; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

[illegible]

RESULT 3  
US-09-782-953-8  
; Sequence 8, Application US/09782953  
; Patent NO. US20020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHMEU, BEVERLY

Query Match	77.28;	Score 720.8;	DB 10;	Length 3184;	
Best Local Similarity	95.98;	Pred. No. 1.4e-226;			
Matches 740;	Conservative	0;	Mismatches 32;	Indels 0;	Gaps 0;
163	QY	CTCGTTGTTTCGGTGCATATGTTACCAAGTTCACGAGTTCAGTCTGTTTGAAGGAGAAAGACGACGAGGAAAA	222		
222	Db				
223	QY	ATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGTGACGTTCAGCTATTTAAGAGTTT	282		
282	Db				
283	QY	CAGAGGTGTCCGTATATAAATTCAGCAATCCTAAATCTCAGCCGAGCTAGGATAGAGCT	342		
342	Db				
343	QY	TCATGAATCCCAATTCAGAGGGAAAAATTAAGCTCTACTTTGCACAGGTTTCAGACTCC	402		
402	Db				
403	QY	AGAGACAGATGGAGACAAACTGGACTTGGCTCCACCCAGCTGCGCAACACGTTTCTCAT	462		
462	Db				
463	QY	CTGCCCCCTTCCTCCCACTCTGTGGCTGGCAGCCCATCAAGATGCCACGCCAGTCT	522		
522	Db				
523	QY	CAACTATGACCTCCTCTATGCTGTGGCCAAACTTAGCACGAGGAGAGAAGTATCAGCTCCA	582		
582	Db				
583	QY	TGCAGGACATGAGTGCACCCCAAGTGTGCTGTCGACGTGTGGACAGTGCACATAGAGGA	642		
642	Db				
643	QY	AGAAGAGGACCCAAAGACTTCCCAAGCCAAAATCATCCAACTCGGGCTCTCGCCT	702		
702	Db				
703	QY	GCCACCTTCGGTGCACACTGAGTGCCTGCTCTCTCTCGATATAGCCGCTCTCTT	762		
762	Db				
763	QY	ATCATGCTTTTCCCCCTGTGTTGTCAAAAAAATTCGCTTTAAATTCCTGGGTGTT	822		
822	Db				
823	QY	GGTGTGTTTCAGATTCCTTCTTGTGTTATCAAGCCTCTCGGACAAAAGGCTAGGAAAGGT	882		
882	Db				
883	QY	GATATGTCCTCGATCATATCATACCATTAAAGTATACCCATTATTTAGAA	934		
942	Db				
943	QY	GATATGTCCTCGATCATATCATACCATTAAAGTATACCCATTATTTAGAA	993		
	Db				

```

RESULT 2
US-09-782-953-17
; Sequence 17, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTPSD-674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn ver. 2.1

```



```

RESULT 4
US-09-782-953-20
; Sequence 20, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21

```

Query Match	29.98;	Score 279.2;	DB 10;	Length 828;
Best Local Similarity	66.08;	pred. No. 3.5e-81;		
Matches 420;	Conservative 0;	Mismatches 213;	Indels 3;	
QY	85	AGACAGGACTGGGCTGTCACTGGTTGTTTGCAGAGAAGCGCTTTTCAAGCAAA		
Db	85	AGRAGAGGAAGAAGAGATGATTTTGGTGAAATGAAGATGATTTGGATGAGA		
QY	145	CTTCAATAGCTCCCAACTCGTGTGTTTCGCTGCAATGTTACCAAGTCAGTGT		
Db	145	TTTAAAGTGATGCGCTACCTCACTTTTGTCTGACAGCTCCATGAAGCAGCTGT		
QY	205	AGAAGAGCAGCAGGAAAAATTTAGGGACTGTTTCGGACTTATGATGACTGTGTG		
Db	205	ACCAGACAGAGGAAAGATTTGAAGCACTCTTCACCATCTATGATGACCAAGT		
QY	265	CCAGCTATTAAAGATTTTCAGACGTGTCGGTATAAACTTCAGCAATCCTAAATC		
Db	265	TCAGTGTTTAAAGCTTTAGAGAGTCAGAAATTAATTCAGCAAACTCGAAGC		
QY	325	CCGAGCTAGGATAGAGCTTCATGAACCCAAATTCAGAGGGAATAAATTAAGCTC		
Db	325	AGACGCCAATAGAAACTCCACGAAACAGACTTCAATGGGCAGAAGCTAAAGCT		
QY	385	TGCACAGGTTCCAGCTCCAGACAGAGATGGAGACAAACTGCATGCGCTCCACC		
Db	385	TGCACAGGTCAGATGTCGGCGAAGTCGGGACAGTCTATCTCTCGCCGCGCC		
QY	445	TGCCAAACAGTTCTCATCTCGCCCCCTTCTCCCGACCTGTGGCTGGCAGCGCC		
Db	445	TGTCACGAGTTCCTCATCTCCCTCCAGCTCTCCCGACCTGCGGGTGGGAAGCA		
QY	505	CGATGCCACGCCAGTCTCTCAACTATGACCTCCTCTATGCTGTGGCCAAACTAGG		
Db	505	AGATGCGATGCGCTGTTATAAATTTATGATTTACTCTGTGCTGTTTCCAAATGGG		
QY	565	AGAAAGTATGAGCTCCATCAGGGACTGAGTCCACCCCAAGTCGTGCTGCACAA		
Db	565	AGAAATATGAATTTTACGCGGGAACAGAGTCGACACCCAGCGTGGTGTTCAT		
QY	625	CGACGTGACATAGAGGAAGAAGAGGACCCCAAGACTTCCCAAGAGCCAAATATG		
Db	625	TGAAGTGAAACTGAAGAGGAAGAAGAGACAAAA--ACCCCAACAGAAAAATTT		
QY	685	AACTCGGGCTCTGGGCTGCCACCTTCGCTGTCCAA 720		
Db	682	GACAAGCGCCCGACCTTCGACCGCAGCTGTAA 717		

RESULT 5  
US-09-782-953-23  
; Sequence 23, Application US/09782953  
; Patent No US2020150953A1  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, R. SANDERS  
; APPLICANT: ROTHERMEL, BEVERLY  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR  
; TITLE OF INVENTION: CALCINEURIN INHIBITORS



; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 597

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(594)

US-09-782-953-5

Query Match 22.4%; Score 209.6; DB 10; Length 597;  
Best Local Similarity 64.8%; Pred. No. 2.7e-58;  
Matches 333; Conservative 0; Mismatches 169; Indels 12; Gaps 1;

QY 138 TCACGTGACTTCAATGACCTCCCAACTCGTTGCTGCTGCAATGTTCCAGTCAGTGT 197  
DB 8 TTAGGAGACTTTAGCTACAAATTTAGCTCCCTGATTTGCTGTGTCGCAACGATGATGCT 67  
QY 198 TTGAGGAG 127  
DB 68 TCAGGAAAGTGAG 257  
QY 258 TGACGCTCCAGCTATTAAAGAGTTTTCAGAGCTGTTTTCGAGCTTATGATGACTGTG 127  
DB 68 TCAGGAAAGTGAG 257  
QY 128 CCACCTCCAGCTATTAAAGAGTTTTCAGAGCTGTTTTCGAGCTTATGATGACTGTG 317  
DB 188 CTGACGCGGATGCGAGCTGCGGCTGCAAGACCGAGTTCTTGGGAGAGAAATGAAGT 247  
QY 378 TCTACTTTGACAGAGTTTCAGACTCCAGAGACAGATGAGAGAGAGAGAGAGAGAGAG 437  
DB 248 TGTATTTTGTCTCAGACT-----TTACACATAGGAAGTTTTCACACCTGGCTCCGC 295  
QY 438 CCAGAGCTGCAACAGAGTTTCTCATCTCGCCCTTCTCCCTCCCTCCCTCCCTCCCTCC 497  
DB 296 CCAATCCCGCAACAGAGTTTCTCATCTCGCCCTTCTCCCTCCCTCCCTCCCTCCCTCC 355  
QY 498 CCATCAACGATGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAG 557  
DB 356 AAGTAGAAGATGCCACCCCGTCATATGATGATGATGATGATGATGATGATGATGATGAT 415  
QY 558 GACGAGGAGAGATGAGTCCATGCGAGGAGTCCATGCGAGGAGTCCATGCGAGGAGTCC 617  
DB 416 GCGCAGGAGAGAGATGAGTCCATGCGAGGAGTCCATGCGAGGAGTCCATGCGAGGAGT 475  
QY 618 ACGTGTGCGACAGTGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 509  
DB 476 ACGTGTGCGACAGTGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

## RESULT 8

US-09-782-953-1

; Sequence 1, Application US/09782953

; Patent No. US20020150953A1

; GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS

; APPLICANT: ROTHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; FILE REFERENCE: US/09/782,953

; CURRENT APPLICATION NUMBER: 60/216,601

; PRIOR FILING DATE: 2001-02-13

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 599

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-782-953-1

## Query Match

Best Local Similarity 21.2%; Score 197.6; DB 10; Length 599;  
Matches 332; Conservative 0; Mismatches 169; Indels 13; Gaps 2;

QY 138 TCACGTGACTTCAATGACCTCCCAACTCGTTGCTGCTGCAATGTTCCAGTCAGTGT 197  
DB 8 TTAGGAGACTTTAGCTACAAATTTAGCTCCCTGATTTGCTGTGTCGCAACGATGATGCT 146  
QY 198 TTGAGGAG 257  
DB 147 TCAGGAAAGTGAG 206  
QY 258 TGACGCTCCAGCTATTAAAGAGTTTTCAGAGCTGTTTTCGAGCTTATGATGACTGTG 317  
DB 207 CCACCTCCAGCTATTAAAGAGTTTTCAGAGCTGTTTTCGAGCTTATGATGACTGTG 266  
QY 318 CTGACGCGGATGCGAGCTGCGGCTGCAAGACCGAGTTCTTGGGAGAGAAATGAAGT 377  
DB 267 CTGACGCGGATGCGAGCTGCGGCTGCAAGACCGAGTTCTTGGGAGAGAAATGAAGT 326  
QY 378 TCTACTTTGACAGAGTTTCAGACTCCAGAGACAGATGAGAGAGAGAGAGAGAGAGAG 437  
DB 327 TGTATTTTGTCTCAGACT-----TTACACATAGGAAGTTTTCACACCTGGCTCCAC 373  
QY 438 CCAGAGCTGCAACAGAGTTTCTCATCTCGCCCTTCTCCCTCCCTCCCTCCCTCCCTCC 497  
DB 374 CCAATCCCGCAACAGAGTTTCTCATCTCGCCCTTCTCCCTCCCTCCCTCCCTCCCTCC 433  
QY 498 CCATCAACGATGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAG 557  
DB 434 AAGTAGAAGATGCCACCCCGTCATATGATGATGATGATGATGATGATGATGATGATGAT 493  
QY 558 GACGAGGAGAGATGAGTCCATGCGAGGAGTCCATGCGAGGAGTCCATGCGAGGAGTCC 617  
DB 494 GCGCAGGAGAGAGATGAGTCCATGCGAGGAGTCCATGCGAGGAGTCCATGCGAGGAGT 553  
QY 618 ACGTGTGCGACAGTGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 551  
DB 554 ACGTGTGCGACAGTGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

## RESULT 9

US-09-782-953-14

; Sequence 14, Application US/09782953

; Patent No. US20020150953A1

; GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS

; APPLICANT: ROTHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; FILE REFERENCE: US/09/782,953

; CURRENT APPLICATION NUMBER: 60/216,601

; PRIOR FILING DATE: 2001-02-13

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 2212

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (25)..(615)

US-09-782-953-14

## Query Match

Best Local Similarity 20.8%; Score 194.4; DB 10; Length 2212;  
Matches 365; Conservative 0; Mismatches 211; Indels 24; Gaps 2;

QY 143 GACTTCAATGACCTCCCAACTCGTTGCTGCTGCAATGTTCCAGTCAGTGTGAA 202  
DB 37 GACCTGAGAGACCTGCCAGCGCCACCATCGCTGTACCTGGAGACCCGCGTGTCTGTG 96

217	AAAGTGAACACGAGGCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACT	276
264	TCCAGGTATTTAAGAGTTTCAGACGTGTCGGGTATAAATTCAGCAATCTTAATCTGCAG	323
277	TTCAAGTATTTAAGAGCTTCAACAGAGTCAAGATAAATTCAGCAATCTTAATCTGCAG	336
324	CCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAGAGGGAAAAATTAAGCTCTACT	383
337	CAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAAATGAAGTTATAT	396
384	TTGCACAGGTTTCAGACTCCAGACAGATGGAGACAAATGCACTTGGGTCCACCCACG	443
397	TTGCTCAG-----ACCTTACATATAGAGCTCACACCTGGCTCCGCCAAATC	444
444	CTGCCAAACAGTTTCTCATCTCGCCCTCTCTCCACACCTGTTGGTGGCAGCCCCATCA	503
445	CAGACAGCAGTTTCTGATCTCCCTCTCCGCTCTCCGAGTGGATGGAAACAAGTGG	504
504	ACGATGCCACGCCAGTCTCTCAACTATGACCTCTCTATGTGTGGCCAAATAGGACCG	563
505	AAGATGCCAGCCCGAGTCAATAAATGATGATCTCTTATATGCCATCTCCAAGCTGGGGCAG	564
564	GAGAGAACTGATGAGCTCCATGCAGGAGTCTGAGTCCACCCCAAGTGTGCTGCACGTG	623
565	GGGAAAGTATGAATTCGACGAGCTGACACCACTCCGAGCGTGGTGTCCATGTAT	624
624	CGCAGCTGACATAGAC-----GAAGAGAGGAGCCCAAAAGACTTCCCCCAAGC	671
625	GTGAGAGTATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	684
672	CAAAATATCAAACTCGGCGCTCTGCGCTGCCACCTCGTGTCCAACTGACGTGCC	730
685	CAAAATATCCAGACGAGGAGGCGGAGTACACGCCGATCCACCTGAGCTGAACTGCG	743

RESULT 11  
 US-10-044-090-255  
 ; Sequence 255, Application US/10044090  
 ; Publication No. US20020137081A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Olga Bandman  
 ; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATED BY  
 ; FILE REFERENCE: PA-0028 US  
 ; CURRENT APPLICATION NUMBER: US/10/044,090  
 ; CURRENT FILING DATE: 2002-01-09  
 ; NUMBER OF SEQ ID NOS: 850  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 255  
 ; LENGTH: 2355  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No. US20020137081A1 5511889CB1  
 US-10-044-090-255

Query Match	20.0%;	Score 187;	DB 13;	Length 2355;
Best Local Similarity	60.1%;	Pred. No. 1.8e-50;		
Matches 360;	Conservative	0;	Mismatches 215;	Indels 24;
Gaps				

QY	144	ACTTCAATGACCTCCGCCAACTCGTGTGGTGGTCAATGTTCCACGACGTGTTGAAG	203
Db	164	ACTTAACTACAGTTTGTAGTCCCTGATTTGGTGGTGGTCAATGATATCTTCACGC	223
QY	204	GAGAGAGGAGCAAGGAAATTTGAGGGACTGTTTCGGACCTTATGATGACGTGTGACGT	263
Db	224	AAAGTGAACACGAGGCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACT	283
QY	264	TCCAGTATTTAAGAGTTTTCAGACGTTGTCGGGTATTAACCTTCAGCAATCTTAATCTGCAG	323
Db	284	TTCAAGTATTTAAGAGCTTCAACACGAGTCAAGATAAATCTTACGAAACCCCTCTCCGAG	343
QY	324	CCGAGCTAGGATAGAGCTTCAATGAACCCCAATTCAGAGGGAAAAATTAAGCTCTACT	383

QY	203	GGAGAAGAGAGCAAGGAAAAAATTTGAGGACATGTTTCGGACTTATGATGACTGTGTGACG	262
Db	97	GACGCCGTGCGCGGCCAAATTTGAGTCCCTCTTTAGGACGATATGCAAGGACATCAC	156
QY	263	TTCAGCTATTTAAGAGTTTTCAGACGTGTCGGTATAAACTTCAGCAAACTCTAAATCTCCA	322
Db	157	TTTCAGTATTTTAAAGAGCTTCAACGAGTCAGATAAACTTCAGCAAACTCTCTCCGCA	216
QY	323	GCCGAGCTAGGATAGAGCTTCATGAAACCAATTCAGAGGGAAAAAATTAAGCTCTAC	382
Db	217	CGAGATCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAGGAAATGAAGTTATAT	276
QY	383	TTTGACACAGTTTCAGACTCCAGAGACAGATGGAGAAAACTGCACATTTGGCTCCACCCCA	442
Db	277	TTTGCTCAG-----ACCTTACATAGGAAGCTCACACCTGGCTCGGCCAAAT	324
QY	443	CCTGCCAAACAGTTTCTCATCTCGCCCCCTTCCTCCCACTGTTGGCTGGGAGCCCATC	502
Db	325	CCAGACAAGCAGTTTCTCATCTCCCTCCCGCTCTCCGACGATGGGATGGAAACAAGTG	384
QY	503	ACGATGCCAGCCAGTCTCCTCAACTATGACCTCCTCTATGCTGTGGCCAACTAGGACCA	562
Db	385	GAGATGGGACCCAGTGCATAACTATGATCTCTTATATGCCATCTCAAAGCTGGGGCCA	444
QY	563	GGAGAGAAGTATGAGCTCCATCGAGGAGCTGAGTCCACCCCAAGTGTCTGTCGACGTG	622
Db	445	GGGGAAGATATGAATTGCACGACGAGTGACACCACTCCACGCGTGGTGCATGTA	504
QY	623	TGGCAGACTGACATAGAG-----GAAGAAGAGGACCCAAAAGACTTCCCCAAAG	670
Db	505	TGTGAGAGTGATCAAGGAGAGGAGGAGGAGGAATGGAAGAATGAGGAGACCTAAG	564
QY	671	CCAAAATCATCCAACTCGGGCTCTGGCCTGCCACCTCCGTTGCCACTGAGCTGCC	730
Db	565	CCAAAATATCCAGACGAGGCGGGAGTACGCGCGATCCACCTCAGCTGACCTGGC	624

RESULT 10

US-09-782-953-11

; Sequence 11, Application US/09782953

; Patent No. US20020150953A1

; GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS

; APPLICANT: ROPHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)

; FILE REFERENCE: UTSD:674P21

; CURRENT APPLICATION NUMBER: US/09/782,953

; CURRENT FILING DATE: 2001-02-13

; PRIOR APPLICATION NUMBER: 60/216,601

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 11

; LENGTH: 2331

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (144)..(734)

US-09-782-953-11

Query Match 20.0%; Score 187; DB 10; Length 2331;

Best Local Similarity 60.1%; Pred. No. 1.8e-50;

Matches 360; Conservative 0; Mismatches 215; Indels 24; Gaps 2;

QY	144	ACTTCAATGACCTCCCAACTCGTTGTTTGGCTGCAATGTTCCACAGTCAGTGTGTTGAAG	203
Db	157	ACTTTAACTACAGCTTTTACTGCTCCCTGATGTCCTGTGTGGCAAAACAGTATATCTTCAGCG	216
QY	204	GAGAAGAGAGCAAGGAAAAATTTGAGGAGTCTTTTCGGACTTATGATGACTGTGTGACGT	263

Db 344 CAGATGCCAGGCTCCAGCTGCATAGACTGTTCTGGGAAAGGAAATGAAGTTATATT 403  
 QY 384 TTGACAGGTTTCAGACTCCAGAGACAGATGGAGACAAACTGCACCTGGCTCCACCCAGC 443  
 Db 404 TTGCTCAG-----ACCTTACACATAGGAAGCTCACACCTGGCTCCGCAATC 451  
 QY 444 CTGCCAAACAGTTTCTCATCTCGCCCTTCTCCACACCTGGTGGTGAGAGCCCATCA 503  
 Db 452 CAGACAAGCAGTTTCTGATCTCCCTTCCGCGCTCTCCGCGAGTGGATGGAACAAGTG 511  
 QY 504 ACGATGCCAGCCAGTCTCAACTATGACCTCTCTATGCTGTGGCCAAACTAGGACAG 563  
 Db 512 AAGATGGGACCCAGTGCATTAATGATCTCTTATGCTTCCAGCTCCAGCTGGGCGAG 571  
 QY 564 GAGAGAAGTATGAGCTCCATCCAGGAGTGCAGTCCACCCCAAGTGTGCTGCAGTGT 623  
 Db 572 GGGAAAGTATGATTCACGCGAGCTGACACCACTCCAGCGTGTGCTCATGTAT 631  
 QY 624 GCGACAGTGACATAGAG-----GAAGAAGAGGACCCCAAGACTTCCCCAAAGC 671  
 Db 632 GTGAGAGTGATCAAGAGAAGGAGGAGAGAGGAATGGAAGAATGAGAGACCTAAGC 691  
 QY 672 CAAAATCATCCAACTCGCGCTCGGCTGCGCCCTCCGCTGCTCCAACTGAGCTGCC 730  
 Db 692 CAAAATTCACAGCAGGAGCGGAGTACACGCGGATCCACCTCAGCTGAAGTGGC 750

## RESULT 12

US-10-084-817-36  
 ; Sequence 36, Application US/10084817  
 ; Publication No. US20030119009A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Susan Stuart  
 ; APPLICANT: Jed G. Nuchtern  
 ; APPLICANT: Sharon E. Plon  
 ; APPLICANT: Jason M. Shohet  
 ; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
 ; FILE REFERENCE: PA-0046 US  
 ; CURRENT APPLICATION NUMBER: US/10/084,817  
 ; PRIOR FILING DATE: 2002-02-25  
 ; PRIOR APPLICATION NUMBER: 60/270,784  
 ; PRIOR FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 365  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 36  
 ; LENGTH: 2355  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No. US20030119009A1 5511889CB1  
 US-10-084-817-36

Query Match 20.0%; Score 187; DB 14; Length 2355;  
 Best Local Similarity 60.1%; Pred. No. 1.8e-50;  
 Matches 360; Conservative 0; Mismatches 215; Indels 24; Gaps 2;  
 QY 144 ACTTCAATGACCTCCCACTCGTGTGTTGCGTGCATGTTTCCAGTCAAGTGTGAAG 203  
 Db 164 ACTTCACTACAGTTTGTAGTCCCTGATTCCTGTGTCGCAACAGTGTATCTTCAGCG 223  
 QY 204 GAGAGAGACCAAGGAAAAATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGTACGT 263  
 Db 224 AAGTGAACACGAGGCGCAATTTGAGTCCCTCTTTAGGAGCTGATGACAGGACATCACT 283  
 QY 264 TCCAGCTATTAAAGAGTTTCAGAGCTGTCGGTATAACTTACAGTAATCTTAATCTGCAG 323  
 Db 284 TTCAGTATTTTAAAGAGCTTCAACGAGTCAAGTAATTAATTCAGCAACCCCTTCTCCGCG 343  
 QY 324 CCGGAGCTAGGAGTTCATGAACCCAAATTCAGAGGGGAAAAATTAAGAGCTCTACT 383  
 Db 344 CAGATGCCAGGCTCCAGCTGCATAGACTGAGTTCTGGGAAAGGAAATGAAGTTATATT 403

QY 384 TTGACAGGTTTCAGACTCCAGAGACAGATGGAGACAAACTGCACCTGGCTCCACCCAGC 443  
 Db 404 TTGCTCAG-----ACCTTACACATAGGAAGCTCACACCTGGCTCCGCAATC 451  
 QY 444 CTGCCAAACAGTTTCTCATCTCGCCCTTCTCCACACCTGGTGGTGAGAGCCCATCA 503  
 Db 452 CAGACAAGCAGTTTCTGATCTCCCTTCCGCGCTCTCCGCGAGTGGATGGAACAAGTG 511  
 QY 504 ACGATGCCAGCCAGTCTCAACTATGACCTCTCTATGCTGTGGCCAAACTAGGACAG 563  
 Db 512 AAGATGGGACCCAGTGCATTAATGATCTCTTATGCTTCCAGCTCCAGCTGGGCGAG 571  
 QY 564 GAGAGAAGTATGAGCTCCATCCAGGAGTGCAGTCCACCCCAAGTGTGCTGCAGTGT 623  
 Db 572 GGGAAAGTATGATTCACGCGAGCTGACACCACTCCAGCGTGTGCTCATGTAT 631  
 QY 624 GCGACAGTGACATAGAG-----GAAGAAGAGGACCCCAAGACTTCCCCAAAGC 671  
 Db 632 GTGAGAGTGATCAAGAGAAGGAGGAGAGGAATGGAAGAATGAGAGACCTAAGC 691  
 QY 672 CAAAATCATCCAACTCGCGCTCGGCTGCGCCCTCCGCTGCTCCAACTGAGCTGCC 730  
 Db 692 CAAAATTCACAGCAGGAGCGGAGTACACGCGGATCCACCTCAGCTGAAGTGGC 750

## RESULT 13

US-09-925-302-347  
 ; Sequence 347, Application US/09925302  
 ; Patent No. US20020044941A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA104  
 ; CURRENT APPLICATION NUMBER: US/09/925,302  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05918  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 896  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 347  
 ; LENGTH: 2358  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-925-302-347

Query Match 19.4%; Score 181.6; DB 9; Length 2358;  
 Best Local Similarity 60.3%; Pred. No. 1.1e-48;  
 Matches 362; Conservative 2; Mismatches 211; Indels 25; Gaps 3;  
 QY 143 GACTTCAATGACCTCCCACTCGTGTGTTGCGTGCATGTTTCCAGTCAAGTGTGA 202  
 Db 126 GACCTGCAGAGCTGCCAGCGCCACCATCGCTGTCACCTGGACCGCGCTGTCTGTG 185  
 QY 203 GCGAAGAGAGCAAGAAAAATTTGAGGAGCTGTTTCGGAGCTTATGATGACTGTGTGAGC 262  
 Db 186 GACGCGCTGTGCGGCGCAAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACC 245  
 QY 263 TTCCAGCTATTTAAGAGTTTTCAGAGCTGTCCGTATATAAACTTCAGCAATCTTAAATCTGCA 322  
 Db 246 TTTCAGTATTTTAAAGAGCTTCAACGAGTCAAGATAAACTTCAGCAACCCCTTCTCCGCA 305  
 QY 323 GCCCAGCTAGGATAGACTTCATGAACCCCAATTCAGAGGGGAAAAATTAAGAGCTCTAC 382  
 Db 306 CGAGATGCCAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAATGAAGTTATAT 365  
 QY 383 TTTCACAGGTTTCAGACTTCCAGAGACAGATGGAGACAACTGCAGTGTGGCTCCACCCAG 442  
 Db 366 TTGTCTCAG-----ACCTTACACATAGGAAGCTCACACCTGGCTCC-GCAAT 412  
 QY 443 CCTGCCAAACAGTTTTCATGCTGCCCCCTTCTCCACCTGTTGGCTGCGACGCCATC 502

Tue Aug 26 09:11:05 2003

```

344 CAGTCATAAACTATGATCTCTTATATGCCATCTCTCAAGCTGGGGCCAGGGGAAAGATG 403
576 AGTCCATGAGGAGGAGTGTGACCCCAAGTGTGCTGCGTGTGCGACAGTGTGCGACAGTGA 635
404 AATGTCAGCGAGGAGTGTGACCCCAAGTGTGCTGCGTGTGCGACAGTGTGCGACAGTGA 463
636 TAGAG-----GAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 683
464 AAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 523
684 AAATCGGGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 730
524 AGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 570

```

```

RESULT 15
US-10-102-524-1746
; Sequence 1746, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Mannion, Paul A.
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1746
; LENGTH: 2173
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-1746

```

```

Query Match 19.0%; Score 177.4; DB 14; Length 2173;
Best Local Similarity 62.0%; Pred. No. 2.5e-47;
Matches 327; Conservative 0; Mismatches 176; Indels 24; Gaps 2;

```

```

; QY 216 AGGAAAAATTTGAGGAGTCTTTGCGGACTTATGATGACTGTGTGACGCTTCCAGCTATTTA 275
; DB 56 ATGCCAAATTTGAGTCTCTTTAGGAGGATGATGAGGAGGATGATGAGGAGGATGATG 115
; QY 276 AGAGTTTCAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 335
; DB 116 AGAGCTTCAACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 175
; QY 336 TAGAGCTTCATGAAACCCCAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 395
; DB 176 TCCAGCTGCATAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 231
; QY 396 AGACTCCAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 455
; DB 232 -----ACCTTACATAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 283
; QY 456 TTCTCATCTCCGCTCCCTTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 515
; DB 284 TTCTCATCTCCGCTCCCTTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 343
; QY 516 CAGTCCCACTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 575
; DB 344 CAGTCCCACTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 403
; QY 576 AGCTCCATGCGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 635
; DB 404 AATTCAGCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 463
; QY 636 TAGAG-----GAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 683
; DB 464 AAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 523

```

```

413 CCAGCAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCTGAGTGGGATGAAACAAGT 472
503 AAGCATGCCACGCGAGTCTCAACTATGAGTCTCTTATGCTGTGCGCAACTAGNACCA 562
473 GAAGATCGGAGCCGAGTCAACTATGATCTCTTATGCTGTGCGCAACTAGNACCA 532
563 GGAGAGAACTATGAGTCTCAACTATGAGTCTCTTATGCTGTGCGCAACTAGNACCA 622
533 GGGGAAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 592
623 TGCCACAGTGCATAGAG-----GAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 670
593 TGTRAGATGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 652
671 CCAAAATCATCAAACTCGGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 730
653 CCAAAATCATCAAACTCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 712

```

```

RESULT 14
US-09-880-107-3340
; Sequence 3340, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2000-06-14
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3340
; LENGTH: 2173
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U28833
US-09-880-107-3340

```

```

Query Match 19.0%; Score 177.4; DB 10; Length 2173;
Best Local Similarity 62.0%; Pred. No. 2.5e-47;
Matches 327; Conservative 0; Mismatches 176; Indels 24; Gaps 2;

```

```

; QY 216 AGGAAAAATTTGAGGAGTCTTTGCGGACTTATGATGACTGTGTGACGCTTCCAGCTATTTA 275
; DB 56 ATGCCAAATTTGAGTCTCTTTAGGAGGATGATGAGGAGGATGATGAGGAGGATGATG 115
; QY 276 AGAGTTTCAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 335
; DB 116 AGAGCTTCAACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 175
; QY 336 TAGAGCTTCATGAAACCCCAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 395
; DB 176 TCCAGCTGCATAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 231
; QY 396 AGACTCCAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 455
; DB 232 -----ACCTTACATAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 283
; QY 456 TTCTCATCTCCGCTCCCTTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 515
; DB 284 TTCTCATCTCCGCTCCCTTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 343
; QY 516 CAGTCCCACTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 575

```

Qy 684 AACTCGCGCTCCTGGCCCTGCCACCCCTCCGTGTCTCAACTGAGCTGCC 730  
Db 524 AGACCGAGGCGCGGAGTACACGCCGATCCACCTCAGCTGAAGTGGC 570

Search completed: August 24, 2003, 13:24:00  
Job time : 326 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2003, 09:53:05 ; Search time 2647 Seconds  
(without alignments)  
8575.890 Million cell updates/sec

Title: US-10-030-613-3  
Perfect score: 934  
Sequence: 1 agaatcattctatcgga.....gtataaccattatttagaa 934

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_plp:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_man:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	705.2	75.5	753	10 AW957141	AW957141 EST369211
2	610.6	65.4	1201	9 AL532776	AL532776 AL532776
3	597.2	63.9	925	13 BQ720798	BQ720798 AGENCOURT
4	589.8	63.1	1055	12 BM925558	BM925558 AGENCOURT

5	581.6	62.3	1201	13 BX420406	BX420406 BX420406
6	531.8	56.9	557	9 AI796134	AI796134 wh42g09.x
7	531.2	56.9	843	14 CA988986	CA988986 AGENCOURT
8	516	55.2	516	9 AL693037	AL693037 DKF2P313K
9	514.6	55.1	2094	11 AK049006	AK049006 Mus muscu
10	513.8	55.0	525	10 AW864635	AW864635 PM4-SN001
11	494.6	53.0	661	14 CB531122	CB531122 742372 MA
12	493.4	52.8	1007	10 BF534085	BF534085 602075073
13	492.6	52.7	530	9 AA814235	AA814235 OD25H08.s
14	480	51.4	749	13 BU847179	BU847179 AGENCOURT
15	478	51.2	899	13 BU503652	BU503652 AGENCOURT
16	474.4	50.8	873	13 BU285474	BU285474 603602128
17	466.2	49.9	847	12 B1545609	B1545609 603187809
18	460.8	49.3	538	14 CB457344	CB457344 714687 MA
19	459.6	49.2	695	14 CB461803	CB461803 721815 MA
20	440.4	47.2	596	14 CB538249	CB538249 776215 MA
21	435.2	46.6	550	12 BG880881	BG880881 1b50f01.y
22	421.8	45.2	485	9 AW655339	AW655339 h189d01.x
23	421.4	45.1	688	10 BG704949	BG704949 602688927
24	420.6	45.0	729	13 BU313385	BU313385 603540473
25	420.4	45.0	842	14 CB575637	CB575637 AGENCOURT
26	414	44.3	758	10 BG709000	BG709000 602675149
27	403.2	43.2	420	10 BF087797	BF087797 IL3-HT061
28	403	43.1	474	9 AW470936	AW470936 xz78c08.x
29	401	42.9	742	10 BG721994	BG721994 602698760
30	395.6	42.4	676	10 BB635249	BB635249 BB635249
31	388.4	41.6	495	10 BF443461	BF443461 260951 MA
32	388.4	41.6	748	12 BM964376	BM964376 UI-M-BQ0-
33	365.4	39.1	433	9 AA248590	AA248590 csh0570.s
34	359.4	38.5	660	9 AL892763	AL892763 AL892763
35	358.4	38.4	681	29 AG056729	AG056729 Pan trogl
36	356.8	38.2	726	29 AG059208	AG059208 Pan trogl
37	353.4	37.8	587	14 CB585503	CB585503 AMGNNUC:N
38	349.4	37.4	626	10 BG701025	BG701025 602682030
39	347.8	37.2	692	14 CB468653	CB468653 734500 MA
40	346	37.0	536	14 CA894406	CA894406 B0185C10-
41	343.8	36.8	633	12 BM489970	BM489970 pqp2n.pk0
42	336	36.0	338	14 T09144	T09144 EST07037.in
43	334.4	35.8	767	13 BU611158	BU611158 UI-M-FC0-
44	331.8	35.5	539	12 BG795691	BG795691 UTSW.SM36
45	329	35.2	347	12 BM994419	BM994419 UI-H-DH0-

## ALIGNMENTS

RESULT 1	AW957141	753 bp	mRNA	linear	EST 01-JUN-2000
LOCUS	EST369211	MAGE	resequences	MAGD Homo sapiens cDNA, mRNA	sequence.
DEFINITION	AW957141				
ACCESSION	AW957141.1	GI:8146824			
VERSION	EST.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	Heide, P., Qi, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C., Holt, I.E., Saced, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.				
AUTHORS	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray				
TITLE	Unpublished				
JOURNAL	Contact: John Quackenbush				
COMMENT	The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3328 Fax: 301 838 0208 Email: johnqu@tigr.org Plate: 102 Seq primer: Reverse. Location/Qualifiers				



Db 754 AAGAGAGGACCCAAAGACTTCCCAAGCCAAATAATCATCCAACTCGGGCTCTGGCC 813  
 QY 702 TGCCACCCCTCGTGTCCAACTGAGTGCCTCTCTCGATAATAGCCGTCCTCTCTTT 761  
 Db 814 TGCCACCCCTCGTGTCCAACTGAGTGCCTCTCTCGATAATAGCCGTCCTCTCTTT 873  
 QY 762 TATCATGCTTTTCCCTCTGT-TTTGTCAAAAAAATTCCTTTAAATTCCTGGGTGT 820  
 Db 874 TATCATGCTTTTCCCTCTGTAGTTGTCAAAAAAATTCCTTTAAATTCCTGGGTGT 933  
 QY 821 TTGGTGTTCAGATTCTCTCTGTATCAAGCTCTCGGACAAAGAGGCTAGGAAAAG 880  
 Db 934 TTTGGTGTTCAGATTCTCTCTGTATCAAGCTCTCGGACAAAGAGGCTAGGAAAAG 991  
 QY 881 GTGATATGCTCTCTGTATCAATATCAATCCAT 911  
 Db 992 TGTATGCTCTCTGTATCAATATCAATCCAT 1022

RESULT 3  
 BQ720798 925 bp mRNA linear EST 16-JUL-2002  
 LOCUS AGENCOURT\_8230839 Lupski\_sympathetic\_trunk Homo sapiens cDNA clone  
 DEFINITION IMAGE:6191156 5', mRNA sequence.  
 BQ720798  
 VERSION BQ720798.1 GI:21859695  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 925)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgraphs@mail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: Agencourt Bioscience Corporation  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1M13591 row: b column: 21  
 High quality sequence start: 20  
 High quality sequence stop: 702.

## FEATURES

Location/Qualifiers  
 1..925  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6191156"  
 /sex="male"  
 /tissue\_type="sympathetic trunk"  
 /dev\_stage="adult, 16 yr"  
 /lab\_host="DH10B"  
 /clone\_lib="Lupski\_sympathetic\_trunk"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
 Noti; Site\_2: SalI; cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCACGCGTCG-3' and  
 5'-GACTAGTTCATGATCGGAGCGCGCT(15)-3'. Size selected >  
 1 kb for average insert length 1.9 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine); available through Life  
 Technologies."

BASE COUNT 227 a 266 c 211 g 221 t  
 ORIGIN

Query Match 63.9%; Score 597.2; DB 13; Length 925;  
 Best Local Similarity 92.1%; Pred. No. 4.5e-157;

Matches 652; Conservative 0; Mismatches 53; Indels 3; Gaps 2;  
 QY 163 CTGCTTTTTCGTCGAATGTTCCACAGTCAGTGTGTAAGGAGAAGAGCAAGAAAA 222  
 Db 218 CACTCTGGT 277  
 QY 223 ATTTGAGGAGCTGTTTTCGGACTTATGATGACTGTGTGAGCTTCCAGCTATTAAAGATTT 282  
 Db 278 ATTTGAGGAGCTGTTTTCGGACTTATGATGACTGTGTGAGCTTCCAGCTATTAAAGATTT 337  
 QY 283 CAGACGTGTCCGTATTAACCTTCAAGCAATCTTAAGCTTCTACTTTGACAGGTTTCAGAGTCC 342  
 Db 338 CAGACGTGTCCGTATTAACCTTCAAGCAATCTTAAGCTTCTACTTTGACAGGTTTCAGAGTCC 397  
 QY 343 TCATGAACCCCAATTCAGAGGAGGAAAAATTAAGCTTCTACTTTGACAGGTTTCAGAGTCC 402  
 Db 398 TCATGAACCCCAATTCAGAGGAGGAAAAATTAAGCTTCTACTTTGACAGGTTTCAGAGTCC 457  
 QY 403 AGAGACAGATGGAGACAACTGCTTCCACCCAGCCCTGCAACACAGTTTCTCAT 462  
 Db 458 AGAGACAGATGGAGACAACTGCTTCCACCCAGCCCTGCAACACAGTTTCTCAT 517  
 QY 463 CTCGCCCTTCTCTCCACCTGTTGGCTGGCAGCCCATCAACAGATGCCAGCCAGTCTCT 522  
 Db 518 CTCGCCCTTCTCTCCACCTGTTGGCTGGCAGCCCATCAACAGATGCCAGCCAGTCTCT 577  
 QY 523 CAACATGATGACCTCTCTATGCTGTGCGCAAACTAGGACGAGAGAGATGATGACTCCA 582  
 Db 578 CAACATGATGACCTCTCTATGCTGTGCGCAAACTAGGACGAGAGAGATGATGACTCCA 637  
 QY 593 TGCAGGAGCTGAGTCCACCCCAAGTGTCTGTGCGACGTGTGCGACAGTGCATAGAGGA 642  
 Db 638 TGCAGGAGCTGAGTCCACCCCAAGTGTCTGTGCGACGTGTGCGACAGTGCATAGAGGA 697  
 QY 643 AGAAGAGGACCCAAAGACTTCCCAAGCCCAAAATCATCAAACTCGGCTCTCTGGCCT 702  
 Db 698 AGAAGAGGACCCAAAGACTTCCCAAGCCCAAAATCATCAAACTCGGCTCTCTGGCCT 757  
 QY 703 GCCACCTCTGCTGCTCAACTGAGTGTCTCTCTCGATAATAGCCGTCCTCTCTTT 762  
 Db 758 GCCACCTCTGCTGCTCAACTGAGTGTCTCTCTCGATAATAGCCGCTCTCTCTTT 816  
 QY 763 ATCATGCTTTTCCCTCTGTTTGTTCARAAAAAATTCGCTTTAAATTCCTGGGTGTTT 822  
 Db 817 ATCATGCTTTTCCCTCTGTTTGTTCARAAAAAATTCGCTTTAAATTCCTGGGTGTTT 876  
 QY 823 GGTGTTTGAATTCCTTCCTCTGTTTATCAAGCTCTCGGACAAAAAG 868  
 Db 877 GGTGTTTGAATTCCTTCCTCTGTTTATCAAGCTCTCGGACAAAAAG 924

RESULT 4  
 BQ925558 1055 bp mRNA linear EST 12-MAR-2002  
 LOCUS AGENCOURT\_6710265 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5763757  
 DEFINITION 5', mRNA sequence.  
 BQ925558  
 ACCESSION BQ925558.1 GI:19375937  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1055)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgraphs@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation

FEATURES	SOURCE
1. <i>Age</i>	1. <i>Age</i>
2. <i>Gender</i>	2. <i>Gender</i>
3. <i>Marital status</i>	3. <i>Marital status</i>
4. <i>Education</i>	4. <i>Education</i>
5. <i>Income</i>	5. <i>Income</i>
6. <i>Occupation</i>	6. <i>Occupation</i>
7. <i>Religion</i>	7. <i>Religion</i>
8. <i>Political affiliation</i>	8. <i>Political affiliation</i>
9. <i>Health status</i>	9. <i>Health status</i>
10. <i>Family size</i>	10. <i>Family size</i>
11. <i>Home ownership</i>	11. <i>Home ownership</i>
12. <i>Travel frequency</i>	12. <i>Travel frequency</i>
13. <i>Spending habits</i>	13. <i>Spending habits</i>
14. <i>Interests and hobbies</i>	14. <i>Interests and hobbies</i>
15. <i>Volunteer work</i>	15. <i>Volunteer work</i>
16. <i>Charitable contributions</i>	16. <i>Charitable contributions</i>
17. <i>Environmental concerns</i>	17. <i>Environmental concerns</i>
18. <i>Community involvement</i>	18. <i>Community involvement</i>
19. <i>Political participation</i>	19. <i>Political participation</i>
20. <i>Healthcare usage</i>	20. <i>Healthcare usage</i>
21. <i>Insurance coverage</i>	21. <i>Insurance coverage</i>
22. <i>Retirement savings</i>	22. <i>Retirement savings</i>
23. <i>Investment preferences</i>	23. <i>Investment preferences</i>
24. <i>Real estate holdings</i>	24. <i>Real estate holdings</i>
25. <i>Art collection</i>	25. <i>Art collection</i>
26. <i>Vehicle ownership</i>	26. <i>Vehicle ownership</i>
27. <i>Travel preferences</i>	27. <i>Travel preferences</i>
28. <i>Dining habits</i>	28. <i>Dining habits</i>
29. <i>Shopping preferences</i>	29. <i>Shopping preferences</i>
30. <i>Technology usage</i>	30. <i>Technology usage</i>
31. <i>Media consumption</i>	31. <i>Media consumption</i>
32. <i>Exercise routine</i>	32. <i>Exercise routine</i>
33. <i>Substance use</i>	33. <i>Substance use</i>
34. <i>Mental health status</i>	34. <i>Mental health status</i>
35. <i>Life satisfaction</i>	35. <i>Life satisfaction</i>
36. <i>Work-life balance</i>	36. <i>Work-life balance</i>
37. <i>Parenting style</i>	37. <i>Parenting style</i>
38. <i>Child development</i>	38. <i>Child development</i>
39. <i>Family dynamics</i>	39. <i>Family dynamics</i>
40. <i>Relationship quality</i>	40. <i>Relationship quality</i>
41. <i>Communication style</i>	41. <i>Communication style</i>
42. <i>Conflict resolution</i>	42. <i>Conflict resolution</i>
43. <i>Emotional stability</i>	43. <i>Emotional stability</i>
44. <i>Stress management</i>	44. <i>Stress management</i>
45. <i>Resilience</i>	45. <i>Resilience</i>
46. <i>Adaptability</i>	46. <i>Adaptability</i>
47. <i>Problem-solving skills</i>	47. <i>Problem-solving skills</i>
48. <i>Decision-making style</i>	48. <i>Decision-making style</i>
49. <i>Time management</i>	49. <i>Time management</i>
50. <i>Goal setting</i>	50. <i>Goal setting</i>
51. <i>Self-discipline</i>	51. <i>Self-discipline</i>
52. <i>Perseverance</i>	52. <i>Perseverance</i>
53. <i>Optimism</i>	53. <i>Optimism</i>
54. <i>Gratitude</i>	54. <i>Gratitude</i>
55. <i>Empathy</i>	55. <i>Empathy</i>
56. <i>Compassion</i>	56. <i>Compassion</i>
57. <i>Kindness</i>	57. <i>Kindness</i>
58. <i>Generosity</i>	58. <i>Generosity</i>
59. <i>Forgiveness</i>	59. <i>Forgiveness</i>
60. <i>Patience</i>	60. <i>Patience</i>
61. <i>Humility</i>	61. <i>Humility</i>
62. <i>Modesty</i>	62. <i>Modesty</i>
63. <i>Respectfulness</i>	63. <i>Respectfulness</i>
64. <i>Politeness</i>	64. <i>Politeness</i>
65. <i>Neatness</i>	65. <i>Neatness</i>
66. <i>Organization</i>	66. <i>Organization</i>
67. <i>Attention to detail</i>	67. <i>Attention to detail</i>
68. <i>Accuracy</i>	68. <i>Accuracy</i>
69. <i>Efficiency</i>	69. <i>Efficiency</i>
70. <i>Productivity</i>	70. <i>Productivity</i>
71. <i>Time management</i>	71. <i>Time management</i>
72. <i>Goal setting</i>	72. <i>Goal setting</i>
73. <i>Self-discipline</i>	73. <i>Self-discipline</i>
74. <i>Perseverance</i>	74. <i>Perseverance</i>
75. <i>Optimism</i>	75. <i>Optimism</i>
76. <i>Gratitude</i>	76. <i>Gratitude</i>
77. <i>Empathy</i>	77. <i>Empathy</i>
78. <i>Compassion</i>	78. <i>Compassion</i>
79. <i>Kindness</i>	79. <i>Kindness</i>
80. <i>Generosity</i>	80. <i>Generosity</i>
81. <i>Forgiveness</i>	81. <i>Forgiveness</i>
82. <i>Patience</i>	82. <i>Patience</i>
83. <i>Humility</i>	83. <i>Humility</i>
84. <i>Modesty</i>	84. <i>Modesty</i>
85. <i>Respectfulness</i>	85. <i>Respectfulness</i>
86. <i>Politeness</i>	86. <i>Politeness</i>
87. <i>Neatness</i>	87. <i>Neatness</i>
88. <i>Organization</i>	88. <i>Organization</i>
89. <i>Attention to detail</i>	89. <i>Attention to detail</i>
90. <i>Accuracy</i>	90. <i>Accuracy</i>
91. <i>Efficiency</i>	91. <i>Efficiency</i>
92. <i>Productivity</i>	92. <i>Productivity</i>
93. <i>Time management</i>	93. <i>Time management</i>
94. <i>Goal setting</i>	94. <i>Goal setting</i>
95. <i>Self-discipline</i>	95. <i>Self-discipline</i>
96. <i>Perseverance</i>	96. <i>Perseverance</i>
97. <i>Optimism</i>	97. <i>Optimism</i>
98. <i>Gratitude</i>	98. <i>Gratitude</i>
99. <i>Empathy</i>	99. <i>Empathy</i>
100. <i>Compassion</i>	100. <i>Compassion</i>

BASE COUNT	244 a	283 c	242 g	225 t	61 others
ORIGIN					
Query Match					
Best Local Similarity					
Matches 623; Conservative					
63.1%; Score 589.8; DB 12;					
86.3%; Pred. No. 5.8e-155;					
0; Mismatches 98; Indels 1; Gaps 1;					

163	QY	CTCTGTTGTTGGCTGCAATGTTTCCACAGTCAGTGTGTTGAAGGAGAGAGACGACGAGAA	222
164			
165			
166			
167	Db	CACTCTGGTTGGCTGTGTGGTGGATCTCGAGGTCTTTACCAATCAGGAGGTTTAAGGAAA	283
168			
169			
170	QY	ATTTCGAGGACGTGTTTCGGACATTATGATGACGTGTGTGACGTTCCAGCTATTTTAAGAGTTT	282
171			
172			
173	Db	ATTTCGAGGACGTGTTTCGGACATTATGATGACGTGTGTGACGTTCCAGCTATTTTAAGAGTTT	343
174			
175			
176	QY	CAGACGTGTCGGTATATAACTTCAGCAATCCTTAATCTGCAGCCCGAGCTAGGATAGAGCT	342
177			
178			
179	Db	CAGACGTGTCGGTATATAACTTCAGCAATCCTTAATCTGCAGCCCGAGCTAGGATAGAGCT	403
180			
181			
182	QY	TCATGAACCCCAATTCAGAGGGAAAAATTTAAAGCTCTACTTTTGCAACAGGTTTCAGACTCC	402
183	Db	TCATGAACCCCAATTCAGAGGGAAAAATTTAAAGCTCTACTTTTGCAACAGGTTTCAGACTCC	463
184			
185			
186	QY	AGACAGATGGAGACAAACTGCACTTGGCTGCACCCCGCCCTGCCAAAACAGTTTCTCAT	462
187			
188	Db	AGACAGATGGAGACAAACTGCACTTGGCTGCACCCCGCCCTGCCAAAACAGTTTCTCAT	523
189			
190			
191	QY	CTCGCCCCCTTCCTCCCCACCTGTGGCTGGCAGCCCATCAAGATGCCACGCCAGTCCT	522
192	Db	CTCGCCCCCTTCCTCCCCACCTGTGGCTGGCAGCCCATCAAGATGCCACGCCAGTCCT	583
193			
194			
195	QY	CAACTATGACCTCTCTATGCTGTGGCCAACTAGGACACGAGGAGAAAGTATGAGCTCCA	582
196			
197	Db	CAACTATGACCTCTCTATGCTGTGGCCAACTAGGACACGAGGAGAAAGTATGAGCTCCA	643
198			
199			
200	QY	TGCAGGACGTGAGTCCACCCCAAGTGTGCTGTCGACGTGTGCACAGTGACATAGAGGA	642
201			
202	Db	TGCAGGACGTGAGTCCACCCCAAGTGTGCTGTCGACGTGTGCACAGTGACATAGAGGA	703
203			
204	QY	AGAGAGGACCCCAAGACTTCCCCCAAGGCCAAAAATCATCAAACTCGGCGCTCTGGCCT	702
205	Db	AGAAGAGACCCCAAGACTTCCCCCAAGGCCAAAAATCATCAAACTCGGCGCTCTGGCCT	763
206			
207	QY	-GCCACCCCTCGGTGTCAACTGAGCTGGCTGCTCTCTCATTAATAGCCGCTCTCCTCTT	761
208			
209	Db	GGCCACCCCTCGGTGTCAACTGAGCTGGCTGCTCTCTCTCATTAATAGCCGCTCTCCTCTT	823
210			
211	QY	TATCATGCTTTTTCCTCCCTGTGTTGTCAAAAAAAATTCGCTTTAAATTCCTGGGTGTT	821
212	Db	TATCATGCTTTTTCCTCCCTGTGTTGTCAAAAAAAATTCGCTTTAAATTCCTGGGTGTT	883

ACCESSION BX420406  
VERSION BX420406.1 GI:30765978  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
JOURNAL  
CONTACT Genoscope

```
FEATURES
source
Faraday Avenue Genoscript Sequence ID: 087691
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cloned="CS0DF028YVC24"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/cloned_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 library. This library was not normalized."
```

BASE COUNT ORIGIN	291 a	308 c	308 g	242 t	52 others
Query Match	62.3%	Score	581.6;	DB 13;	Length 1201;
Best Local Similarity	93.4%;	Pred. No.	1.3e-152;		
Matches	637;	Conservative	12;	Mismatches	28;
				Indels	5;
				Gaps	4
QY	1	AGAATCATACTTTCGGAATGAGGAGGCCAGGCGAGCAGGACACGTCCTCTGAAGATGG	60		
DB	330	AGAATHAAATTTATCGGAATGARGAACCMAGGCGAGCAGGACACGTCCTCTGAARATGG	389		
QY	61	AGGACTTTTCTACTGTGCTCCATGACAGGAGCTGGGCTG--TCACTCGTTGTGTTTTGCA	118		
DB	390	AGGAATTTTTCTTATGTGCTGCATAGAAAGGAGCTGGGCTGATCATACATCCTTTGTTTGGCA	449		
QY	119	GAAGAAGCGTTTCAAGCAATCACTGACTTCAATGACCTCCCACTCGTTGTTTCGGTGC	178		
DB	450	AAAGAAGCGCTTCAAGCAATCACTGACTTCAATGACCTCCCACTCGTTGTTTCGGTGC	509		
QY	179	AATGTTTACCAGTCAGTGTGTTGAAGGAGAAGCAAGGAAAAATTTGAGGAGCTGTTT	238		
DB	510	AATGTTTACCAGTCAGTGTGTTGAAGGAGAAGCAAGGAAAAATTTGAGGAGCTGTTT	569		



## FEATURES

AL693037	GI:19617757
AL693037.1	
EST.	
KEYWORDS	Homo sapiens (human)
SOURCE	
ORGANISM	
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
	1 (bases 1 to 516)
REFERENCE	Koehler, K., Beyer, A., Mewes, W., Well, B. and Wiemann, S.
AUTHORS	EST (Koehler, K., Beyer, A., Mewes, W., Well, B. and Wiemann, S.)
TITLE	Unpublished
JOURNAL	Contact: Koehler K
COMMENT	

CONTACT: Koehrer K  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sponsored by BMFZ (Biomedical Research Center at the Charite,  
Berlin/Germany) within the cDNA sequencing consortium of the German  
Genome Project.  
No sl sequence available.  
This clone (DKFZp313K0237) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Potsdam, Germany. Email: clone@rzd.de.

FEATURES	source
Location/Qualifiers	
i. .516	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="DKFZp313K0237"	
/dev_stage="adult"	
/lab_host="DH10B"	
/clone_lib="j13 (synonym: hlcc2)"	
/note="vector: pTriplex2; site_1: sf1A; site_2: sf1B;	
cdna_collection"	
150.5	118

BASE COUNT	133 a	159 c	106 g	118 t
ORIGIN				

Query Match	55.2%	Score 516;	DB 9;	Length 516;
Best Local Similarity	100.0%	Pred. No. 3e-134;		
Gap	0;	Mismatches	0;	Indels
Gap	0;	Mismatches	0;	Indels

306 GCAATCCTAAATCTGCAGCCCGAGCTTAGATAGAGCTTCATGAACCCCAATTCAGAGGGA 360

DG  
I GCAATCTTAAATTCAGAGCCCGACGGCAGCGAAGGAGGTTCTGTGTTT  
366 AAAATTAAAGCTTACTTTGCACAGGTTTCAGACTCCAGACACACATGGAGCAAACCTGC 425

61 AAAAAATTAAAGCTCTACTTTGCACAGGTTACAGATCCAGAGACAGATGGAGACAAACTGC 120

426 ACTTGGTCCACCCAGCCTGCCAACAGTTCTCATCTCGGCCCCCTTCTCCCAACCTG 480

D6 AC110GCTCCACCGCCAGTTCCTCAACTATGACCTCCTCTATGCTG 545

QV TTGGCTGGCAGGCCCATCAACGATGCCACGCCAGTTCCTCAACTATGACCTCCTCTATGCTG 545

\_\_\_\_\_



```
BASE COUNT      528 a      525 c      460 g      581 t
ORIGIN
Query Match      55.1%; Score 514.6; DB 11; Length 2094;
Best Local Similarity 79.8%; Pred. No. 1.2e-133;
Matches 617; Conservative 0; Mismatches 155; Indels 1; Gaps 1;

QY 163 CTGCTGTTTGGCGTGCATCTCCACGATCAGTGTGTTGAAGGAGAGAGAGCAAGGAAA 222
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 232 CACTCTGGTCCCTGTGTGGATGTGGAGGCTTTTACCACATCAGGAGGTTAAGGAAA 291
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 223 ATTGAGGAGACTGTTTCGGACTTATGATGACTGTGTGACGTTCCAGCTATTTAAGAGTTT 282
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 292 ATTGAGGAGACTGTTTCGGAGCTATGATGAATGTGTGACGTTCCAGCTGTTTAAGAGTTT 351
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 283 CAGAGCTGTCGGTATATAACTTCAGCAATCTCAATCTGACGCGGAGCTAGGATAGACT 342
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 352 CCGACGGGTTTTCGATTAATTTTCAGCCATCCCAATTCGACGGCGTCCCGGATAGAGCT 411
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 343 TCATGAACCAATTCAGAGGAGAAAATTAAGACTCTACTTTGCACAGGTTTCAGACTCC 402
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 412 TCATGACACTCAGTTCAGAGGAGAGAGCTAAACTCTACTTCGCCAGGTCCAGACCCC 471
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 403 AGAGACAGATGGAGACAAATTCGACTTGGCTTCCACCCAGCCTGCCAAACAGTTTTCAT 462
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 472 AGAGACAGATGGAGACAAATTCGACTTGGCTTCCACCCAGCCTGCCAAACAGTTTTCAT 531
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 463 CTCGCCCTCTCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 522
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 532 CTCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 591
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 523 CAACATGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 582
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 592 CAACATGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 651
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 583 TGCAGGAGTGTGCTCCACCCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 642
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 652 CGCTGGAACTGAGTCTACACGCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 711
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 643 AGAAGAGACCAAGACTTCCCAAGCCAAAATCATCCAAACTCGGCGCTCTGGGCT 702
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 712 GGAGGAGACCAAGACTTCCCAAGCCAAAATCATCCAAACTCGGCGCTCTGGGCT 771
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 703 GCCACCTCTCTGCTCCAACTGAGC-TGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 761
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 772 GCCACCTCTCTGCTCCAACTGAGCAGCTTCCATGATGATGATGATGATGATGATGATGAT 831
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 762 TATCATGCTTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 821
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 832 TCCCGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 891
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 822 TGGTTGTTGAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 881
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 892 TGGTTGTTGAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 951
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 882 TGATATGCTCTGATCATATCATACCATTAAGTATACCCATTAATTTAGAA 934
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 952 TGATATGCTCTCCCACTCACTCTACCCATTAACCTTAACCTTAACCTTAACCTTAACCTTA 1004
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
AW864635
LOCUS      525 bp      mRNA      linear      EST 22-MAY-2000
DEFINITION PM4-SN0017-310300-002-al2 SN0017 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW864635
VERSION   AW864635.1 GI:7998685
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 525)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
```

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-PM4-SN0017-310300-002-al2&t3=2000-03-31&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 523.

FEATURES  
source

1..525  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="SN0017"  
/note="Organ: stomach\_normal; Vector: puc18; Site\_1: Sma1; Site\_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 140 a 135 c 120 g 130 t

Query Match 55.0%; Score 513.8; DB 10; Length 525;  
Best Local Similarity 98.7%; Pred. No. 1.3e-133;  
Matches 518; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 16 CGGAATGAG 75  
Db 1 CGGAATGAG 60

QY 76 GTGCTGCATAGACAGGAG 135  
Db 61 GTGCTGCATAGACAGGAG 120

QY 136 AATCACTGACTTCAATGACCTCCCAACTCGTGTGTTGGTGCATATGTTTCCAGCTCAGT 195  
Db 121 AATCACTGACTTCAATGACCTCCCAACTCGTGTGTTGGTGCATATGTTTCCAGCTCAGT 180

QY 196 GTTGAAG 255  
Db 181 GTTGAAG 240

QY 256 TGTGAGCTTCCAGCTTATTTAAGAGTTTCAGAGCTGTCCGTATATAAATTCAGCAATCTTAA 315  
Db 241 TGTGAGCTTCCGCTTATTTAAGAGTTTCAGAGCTGTCCGTATATAAATTCAGCAATCTTAA 300

QY 316 ATCTCAG 375  
Db 301 ATCTCAG 360

QY 376 GCTCTACTTTCACAGGTTTCAGCTCCAGAGACAGATGGAGACAACTCACTTGGCTCC 435  
Db 361 GCTCTACTTTCACAGGTTTCAGCTCCAGAGACAGATGGAGACAACTCACTTGGCTCC 420









Search completed: August 24, 2003, 12:11:21  
Job time : 2652 secs